



UNIVERSIDADE FEDERAL DO PARANÁ

CLAUDIA MARINA SCHELLIN BECKER

DIVERSIDADE GENÉTICA DE POPULAÇÕES AMERÍNDIAS GUARANI E  
KAINGANG DO BRASIL BASEADA EM STRs AUTOSSÔMICOS DE  
APLICAÇÃO FORENSE

CURITIBA

2019

UNIVERSIDADE FEDERAL DO PARANÁ

CLAUDIA MARINA SCHELLIN BECKER

DIVERSIDADE GENÉTICA DE POPULAÇÕES AMERÍNDIAS GUARANI E  
KAINGANG DO BRASIL BASEADA EM STRs AUTOSSÔMICOS DE  
APLICAÇÃO FORENSE

Dissertação apresentada como requisito parcial  
à obtenção do grau de Mestre em Genética,  
Programa de Pós Graduação em Genética,  
Setor de Ciências Biológicas, Universidade  
Federal do Paraná.

Orientadora: Profa. Dra. Danielle Malheiros

Co-Orientador: Prof. Dr. Marcelo Malaghini

Co-Orientador: Prof. Dr. Danilo Gardenal Augusto

CURITIBA

2019

Universidade Federal do Paraná. Sistema de Bibliotecas.  
Biblioteca de Ciências Biológicas.  
(Giana Mara Seniski Silva – CRB/9 1406)

Becker, Claudia Marina Schellin

Diversidade genética de populações ameríndias Guarani e Kaingang do Brasil baseada em STRs autossômicos de aplicação forense. / Claudia Marina Schellin Becker. – Curitiba, 2019.

151 p. : il.

Orientador: Danielle Malheiros

Coorientadores : Marcelo Malaghini e Danillo Gardenal Augusto

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética forense 2. Genética da população humana 3. Índios Guarani 4. Índios Kaingang I. Título II. Malheiros, Danielle III. Malaghini, Marcelo IV. Augusto, Danillo Gardenal V. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 576.58



MINISTÉRIO DA EDUCAÇÃO  
SETOR SETOR DE CIÊNCIAS BIOLÓGICAS  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA -  
40001016006P1

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **CLAUDIA MARINA SCHELLIN BECKER** intitulada: **Diversidade genética de populações ameríndias Guarani e Kalingang do Brasil baseada em STRs autossômicos de aplicação forense**, após terem inquirido a autora e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 28 de Março de 2019.

  
DANIELLE MALHEIROS FERREIRA

Presidente da Banca Examinadora (UFPR)

  
CELSO TEIXEIRA MENDES JUNIOR  
Avaliador Externo (FFCLRP-U)

  
MARCIA HOLSBACH BELTRAME  
Avaliador Interno (UFPR)

*“It would not be scientifically justifiable to speak of a match as proof of identity in the absence of underlying data that permit some reasonable estimate of how rare the matching characteristics actually are”. NRC II, p. 192*



## DEDICATÓRIA

Aos meus pais, Annelie von Baranow Schellin e Osvaldo Otto Schellin (*in memorian*), pelo maravilhoso exemplo de vida, pela presença, amor e dedicação incondicionais, por me ensinarem que o conhecimento é a base de tudo e por não terem medido esforços para garantir meu acesso a ele.

A meu marido, Ermelino Franco Becker, meu sustentáculo, pelo seu amor, apoio, compreensão e colaboração constantes, mesmo diante de minhas ausências, e por sempre ter acreditado em mim e nos meus sonhos.

A meus filhos, Melissa e Felipe, por serem meus amores, minha razão e inspiração, por compreenderem minha falta de tempo e por toda a sua cooperação nos momentos em que precisei.

## **AGRADECIMENTOS**

Agradeço à Deus, que me abriu as portas, me concedeu tantos privilégios e conquistas e ainda me deu oportunidade, força e inspiração para a realização do mestrado.

Agradeço à Polícia Científica do Estado do Paraná e à CAPES, pela disponibilização de recursos financeiros.

Agradeço à Gerência de Laboratórios Forenses por ter autorizado a realização desta pesquisa em suas dependências e por ter apoiado este estudo.

Agradeço aos membros da banca examinadora, Prof<sup>a</sup>. Dr<sup>a</sup>. Márcia Holsbach Beltrame, Prof. Dr. Celso Teixeira Mendes Junior e em especial minha orientadora, Prof<sup>a</sup>. Dr<sup>a</sup>. Danielle Malheiros, pelas valiosas contribuições para a finalização deste trabalho.

Agradeço à minha querida orientadora, Prof<sup>a</sup>. Dr<sup>a</sup>. Danielle Malheiros, pela infinita paciência, pelo tempo, conhecimento e experiência dedicados, guiando-me sempre com tanto profissionalismo e carinho, transmitindo segurança e dando condições para que este trabalho fosse realizado.

Agradeço ao coorientador e perito criminal Dr. Marcelo Malaghini, pelo imenso apoio e suporte oferecidos durante a realização de todas as etapas envolvidas no desenvolvimento do mestrado.

Agradeço ao coorientador Prof. Dr. Danilo Gardenal Augusto, por tão prontamente compartilhar seu conhecimento em genética de populações, pelas brilhantes sugestões e pelo seu precioso tempo dedicado a responder minhas dúvidas.

Agradeço à Profa. Dra. Maria Luíza Petzl-Erler, por disponibilizar as amostras das populações indígenas para que este estudo pudesse ser realizado.

Agradeço à Veronica Calonga Solís, por sempre encontrar um tempinho para mim e pela inestimável ajuda, prontidão e simpatia, em muitos e muitos momentos, tanto presenciais quanto à distância.

Agradeço à minha amiga Marianna por ter plantado a semente e ter alimentado em mim o desejo pelo crescimento profissional, pelo incentivo constante, por me ouvir, acreditar em mim e por me instigar a ser uma pessoa melhor.

Agradeço à minha amiga Maria Christina pelo incentivo desde o comecinho, antes de tudo.

Agradeço à Aline, pela companhia, tranquilidade e ajuda nas etapas práticas em meio a todo o caos do processo da mudança e acreditação do laboratório.

Agradeço à Luíza, pela ajuda em etapas práticas e por me ensinar a psicologia necessária para trabalhar com os programas estatísticos.

Agradeço ao meu cunhado e amigo Carlos Bascuñan, pela preciosa ajuda em todas as minhas dificuldades com a informática (que não foram poucas) e à minha cunhada e irmã Sílvia Becker, pela força e compreensão.

Agradeço à minha irmã, Andrea Frazon, que sempre me apoia e torce pelo meu sucesso.

Agradeço aos meus colegas do LGMF, os peritos criminais Marcelo, Marianna, Luciellen, Leonardo, Hemerson, Maria Christina, Anna Carolina, Josiane, Camila, Paula e Jaqueline, e à técnica Madalena, pela amizade, pelo apoio nos momentos difíceis e pela cooperação sempre que se fez necessária.

Agradeço aos meus amigos Monika e Rildo, pelo carinho, pela hospedagem e pelo computador compartilhado numa etapa importante do mestrado.

Agradeço ao Programa de Pós Graduação em Genética da Universidade Federal do Paraná, por todo o conhecimento transmitido.



## RESUMO

Quando perfis genéticos identificados em um vestígio de crime e uma amostra referência geram coincidências entre todas as regiões genômicas analisadas, o valor desta associação deve ser estatisticamente avaliado a fim de se quantificar a força da evidência encontrada. Tais cálculos congregam informações de frequências alélicas obtidas em amostras representativas de diferentes populações, urbanas ou isoladas, devendo ainda incorporar um fator de correção ( $\theta$ ) para a eventualidade de subdivisão populacional. O valor de  $\theta$  é obtido a partir dos índices R ou F, mais precisamente Rst (ou Fst), utilizados para mensurar a diferenciação genética observada entre populações. Seu valor foi padronizado como sendo 0,01 para populações urbanas e 0,03 para populações isoladas. Entretanto, há incongruências entre este valor e aquele calculado especificamente para populações isoladas. A população brasileira carece de dados e bancos de perfis genéticos para populações isoladas, o que impacta as análises estatísticas. O objetivo deste estudo foi estimar frequências alélicas de 21 marcadores autossômicos amplamente utilizados na genética forense e, através destes, caracterizar a diversidade genética e possível estruturação de populações ameríndias Guarani e Kaingang do Paraná e do Mato Grosso do Sul, a fim de fornecer estimativas estatísticas adequadas que incrementem o poder das associações com as evidências de DNA que envolvam estas populações. Para tanto, a amostra foi composta por 319 indivíduos: 121 Kaingang, oriundos de duas localidades: Ivaí (KIV=61) e Rio das Cobras (KRC=60), e 198 Guaranis, oriundos de três localidades: M'byá do Rio das Cobras (GRC=51), Guarani Nandeva do município de Amambaí (GND=71) e Guarani Kaiowá dos municípios de Amambaí e Porto Lindo (GKW=76). Todos os locos estão em equilíbrio de Hardy-Weinberg. Foram identificados, por população ameríndia, os marcadores mais e menos informativos. As populações Guarani apresentaram diferenciação variando de baixa (Rst=0,040,  $p < 10^{-4}$ ) a alta (Rst=0,155,  $p < 10^{-5}$ ) entre si, e de média (Rst=0,056,  $p < 10^{-5}$ ) a alta (Rst=0,161,  $p < 10^{-5}$ ) com relação às populações Kaingang. As duas populações Kaingang apresentaram diferenciação muito baixa entre si (Rst=0,004,  $p = 0,27$ ). Todas as demais análises que avaliam estruturação populacional (AMOVA, PCoA, Structure, agrupamento hierárquico) apontaram para o mesmo resultado, indicando diferença populacional entre Guaranis, mas não entre Kaingangs, de modo que dados de frequências alélicas destes podem ser reunidos com propósito forense. O poder de discriminação combinado (PDC) e o poder de exclusão combinado (PEC) foram calculados para cada população, tendo-se obtidos os menores valores, respectivamente, de 0,99999999999999995 e 0,99995705 (ambos para GRC), indicando que o painel de marcadores é adequado para análises forenses envolvendo estas populações. Considerando-se a heterogeneidade demográfica existente entre ameríndios e a escassez de informações precisas da origem populacional de um vestígio de crime, foi calculado o valor médio de Fst para 43 populações ameríndias incluindo os Guaranis e Kaingangs, tendo-se obtido o valor de 0,03, que passa a ser considerado o valor de  $\theta$ . Assim, confirma-se a adequação do valor sugerido (0,03) para cálculos forenses que envolvam populações indígenas nas américas.

**Palavras-chave:** Guarani, Kaingang, Ameríndios, Brasil, STR, Theta, Forense.

## ABSTRACT

When DNA profile comparisons between a crime scene trace and a reference sample generate full coincidences, the match probability has to be estimated in order to evaluate the strength of the forensic DNA evidence. Allele frequency validated databases, obtained from representative samples of urban and isolated populations are required for the random match possibility estimations. The eventuality of population substructure can be adjusted by the incorporation of a correction factor referred to as theta ( $\theta$ ), conservatively standardized as 0.01 (for urban populations) and 0.03 (for isolated populations). However, inconsistencies have been found between the standardized value and that calculated specifically from smaller and isolated population's data. Despite of its high miscegenation, Brazil lacks of genetic data for isolated populations. The objectives of this study were to estimate allelic frequencies of 21 autosomal markers widely used in forensic genetics and to characterize the genetic diversity and possible structuring of South Brazil's Amerindian populations Guaraní and Kaingang, obtain the appropriate value for  $\theta$  and, therefore, increase the power of associations with DNA evidences involving these populations. The sample consisted of 319 individuals: (1) 121 Kaingang, from two communities, Ivaí (KIV=61) and Rio das Cobras (KRC=60) and (2) 198 Guaranis from three communities, M'byá do Rio das Cobras (GRC = 51), Guaraní Ñandeva (GND = 71) and Guaraní Kaiowá (GKW = 76). All loci met Hardy-Weinberg equilibrium expectations. The most and the less informative markers were identified for each Amerindian population. Between Guaraní populations low ( $R_{st} = 0.040$ ,  $p < 10^{-4}$ ) to high ( $R_{st} = 0.155$ ,  $p < 10^{-5}$ ) differentiation was found. Regarding Guaraní and Kaingang populations, intermediate ( $R_{st} = 0.056$ ,  $p < 10^{-5}$ ) to high ( $R_{st} = 0.161$ ,  $p < 10^{-5}$ ) differentiation was found. The two Kaingang populations showed very low differentiation between each other ( $R_{st} = 0.004$ ,  $p = 0.27$ ), which justifies the union of both genetic data for databases and forensic calculations. The combined power of discrimination (PD) and the combined power of exclusion (PE) were calculated for each population and the lowest values obtained were, respectively, 0,99999999999999995 e 0,99995705 (both for GRC). Considering the demographic heterogeneity of Amerindian populations in general, the  $F_{st}$  mean value (0,03) was calculated regarding 43 different indigenous populations from the Americas, including Guaranis and Kaingangs. This result confirms the adequacy of the suggested value for  $\theta$  for Amerindian populations.

**Keywords:** Guaraní, Kaingang, Brazil, Amerindian, STRs, Theta, Forensics.

## LISTA DE FIGURAS

<b>FIGURA 1</b> – REPRESENTAÇÃO DE UM MARCADOR STR EM TRÊS INDIVÍDUOS DISTINTOS, CADA QUAL APRESENTANDO DIFERENTES NÚMEROS DE REPETIÇÕES DE TRINUCLEOTÍDEOS. ....	21
<b>FIGURA 2</b> - LOCALIZAÇÃO GEOGRÁFICA DAS ÁREAS DE RESERVA ONDE VIVEM AS POPULAÇÕES INDÍGENAS ANALISADAS NESTE ESTUDO. ....	50
<b>FIGURA 3</b> - FREQUÊNCIAS ALÉLICAS DE 21 MARCADORES STR NAS 5 POPULAÇÕES AMERÍNDIAS. ....	58
<b>FIGURA 4</b> - COMPONENTES DE VARIAÇÃO ENTRE POPULAÇÕES E DENTRO DAS POPULAÇÕES, PARA TODOS OS INDÍGENAS GUARANI (GKW, GND E GRC) E KAINGANG (KIV E KRC). ....	70
<b>FIGURA 5</b> - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA ESTIMADO PARA OS INDIVÍDUOS E PARA AS POPULAÇÕES DE AMERÍNDIOS GUARANIS (GKW, GND E GRC) E KAINGANGS (KRC E KIV). ....	73
<b>FIGURA 6</b> - DENDROGRAMA GERADO COM BASE NOS MARCADORES 21 STR, PARA AS POPULAÇÕES INDÍGENAS. ....	74
<b>FIGURA 7</b> - PCOA BASEADA EM 21 STRS DEMONSTRANDO PADRÕES DE VARIAÇÃO GENÉTICA NAS POPULAÇÕES AMERÍNDIAS ESTUDADAS. ....	75
<b>FIGURA 8</b> - NÚMERO MÉDIO DE ALELOS POR POPULAÇÃO AMERÍNDIA E BRASILEIRA MISCIGENADA, CONSIDERANDO O PAINEL DE 21 MARCADORES STR. ....	96
<b>FIGURA 9</b> - COMPONENTES DE VARIAÇÃO ENTRE POPULAÇÕES E DENTRO DAS POPULAÇÕES, CONSIDERANDO-SE OS INDÍGENAS GUARANI E KAINGANG E A POPULAÇÃO MISCIGENADA BRASILEIRA. ....	99
<b>FIGURA 10</b> - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA ESTIMADO PARA OS INDIVÍDUOS E PARA AS POPULAÇÕES DE AMERÍNDIOS GUARANIS (GKW, GND E GRC), KAINGANGS (KRC E KIV) E POPULAÇÕES MISCIGENADAS BRASILEIRAS. ....	103
<b>FIGURA 11</b> - DENDROGRAMA GERADO COM BASE NOS MARCADORES 21 STRs, PARA AS POPULAÇÕES INDÍGENAS E BRASILEIRA MISCIGENADA. ....	104
<b>FIGURA 12</b> - PCOA DAS FREQUÊNCIAS ALÉLICAS DE 21 STRS TESTADOS PARA OS POVOS AMERÍNDIOS GUARANI E KAINGANG E AS POPULAÇÕES DE CINCO REGIÕES BRASILEIRAS. ....	105
<b>FIGURA 13</b> - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA PARA OS INDIVÍDUOS E POPULAÇÕES DE AMERÍNDIOS GUARANIS, KAINGANGS E POPULAÇÕES	

	MISCIGENADAS PARENTAIS NORTE AMERICANAS COM PREDOMINÂNCIA AFROAMERICANA (AA), EUROPÉIA (EUR), HISPÂNICA (HISP) E ASIÁTICA (ASIAN). ....	109
<b>FIGURA 14</b> -	PCOA REALIZADA ENTRE TODOS OS AS 5 POPULAÇÕES AMERÍNDIAS E POPULAÇÕES PARENTAIS, PARA OS 21 MARCADORES STR. ....	110
<b>FIGURA 15</b> -	PCA REALIZADO ENTRE TODOS OS AMERÍNDIOS E POPULAÇÕES PARENTAIS, PARA 6 MARCADORES STR (D3S1358, D8S1179, D21S11, D18S51, D19S539 e TH01). ....	112
<b>FIGURA 16</b> -	DENDROGRAMA OBTIDO COM BASE EM 6 MARCADORES STR, PARA AS POPULAÇÕES AMERÍNDIAS E PARENTAIS. ....	115

## LISTA DE TABELAS

<b>TABELA 1</b> – VALOR DE P REFERENTE AO EQUILÍBRIO DE HARDY-WEINBERG PARA CADA MARCADOR POR POPULAÇÃO INDÍGENA.....	62
<b>TABELA 2</b> - ALELOS EXCLUSIVOS DETECTADOS NA COMPARAÇÃO DAS FREQUÊNCIAS ALÉLICAS DOS AMERÍNDIOS ENTRE SI.....	65
<b>TABELA 3</b> - PARES DE MARCADORES EM DESEQUILÍBRIO DE LIGAÇÃO EM CADA POPULAÇÃO AMERÍNDIA. ....	67
<b>TABELA 4</b> - AMOVA LOCO POR LOCO BASEADO NOS VALORES DE RST CONSIDERANDO AS POPULAÇÕES INDÍGENAS REUNIDAS. ....	70
<b>TABELA 5</b> - VALORES DE DIFERENCIAÇÃO (Rst) PARA AS 5 POPULAÇÕES INDÍGENAS.....	71
<b>TABELA 6</b> - VALORES DE Rst E Ris PARA AS POPULAÇÕES AMERÍNDIAS. ....	72
<b>TABELA 7</b> - VALORES DE DIFERENCIAÇÃO (RST) PARA 4 POPULAÇÕES INDÍGENAS.....	77
<b>TABELA 8</b> – REGIÃO GEOGRÁFICA, NÚMERO DE ALELOS, HETEROZIGOSIDADE ESPERADA E HETEROZIGOSIDADE OBSERVADA DAS POPULAÇÕES GUARANI E KAINGANG....	78
<b>TABELA 9</b> - MARCADORES FORENSES MAIS E MENOS INFORMATIVOS PARA CADA POPULAÇÃO INDÍGENA.....	79
<b>TABELA 10</b> - VALOR DE PODER DE DISCRIMINAÇÃO COMBINADO (PDC) E PODER DE EXCLUSÃO COMBINADO (PEC), ESTIMADOS PARA O CONJUNTO DE 21 MARCADORES STR, POR POPULAÇÃO INDÍGENA.....	80
<b>TABELA 11</b> – Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR. ....	81
<b>TABELA 12</b> - AMOVA LOCO POR LOCO BASEADO NOS VALORES DE RST CONSIDERANDO AS POPULAÇÕES INDÍGENAS E A POPULAÇÃO MISCIGENADA BRASILEIRA REUNIDAS. ....	100
<b>TABELA 13</b> - VALORES DE RST PARA 4 SUBPOPULAÇÕES INDÍGENAS E POPULAÇÃO BRASILEIRA.....	101
<b>TABELA 14</b> – Valores médios de Rst e Fst para as 4 populações indígenas. ....	117
<b>TABELA 15</b> - Fator de Correção ( $\theta$ ) obtido a partir dos valores médios de Fst para as populações ameríndias. ....	120
<b>TABELA 16</b> - Frequências alélicas da população Guarani Kaiowá (GKW). ....	134
<b>TABELA 17</b> - Frequências alélicas da população Guarani Nandeva (GND). ....	135
<b>TABELA 18</b> - Frequências alélicas da população Guarani M'byá (GRC). ....	136
<b>TABELA 19</b> - Frequências alélicas da população Kaingang – Rio Ivaí (KIV). ....	137
<b>TABELA 20</b> - Frequências alélicas da população Kaingang – Rio das Cobras (KRC). ....	139
<b>TABELA 21</b> - Parâmetros forenses estimados por população ameríndia.....	140
<b>TABELA 22</b> - Número de alelos por marcador e por população. ....	141

<b>TABELA 23</b> - Valores de Rst par a par, entre as populações indígenas, para os marcadores TH01, D21S11, D18S51, D7S820, D16S539, SE33, VWA, D2S1338, D19S433 e D10S1248.....	<b>142</b>
<b>TABELA 24</b> - Valores de Rst par a par, entre as populações indígenas e a população brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, vWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 e D12S391. ....	<b>143</b>
<b>TABELA 25</b> - Resultados de Fst par a par entre Guaranis, Kaingangs, ameríndios (América do Norte, América Central e América do Sul) e Parentais (Japão, África do Sul, Portugal, Alemanha e Espanha).....	<b>148</b>
<b>TABELA 26</b> - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS GUARANI E KAINGANG. ....	<b>151</b>
<b>TABELA 27</b> - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS BRASILEIROS (GUARANI E KAINGANG) E A POPULAÇÃO BRASILEIRA MISCIGENADA. ....	<b>152</b>
<b>TABELA 28</b> - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS BRASILEIROS (GUARANI E KAINGANG) E AS POPULAÇÕES PARENTAIS.....	<b>152</b>



## LISTA DE ABREVIATURAS E SIGLAS

CODIS	Combined DNA Index System
DL	Desequilíbrio de Ligação
EHW	Equilíbrio de Hardy Weinberg
FAM	Frequência Alélica Mínima
FBI	Federal Boureau of Investigation (EUA)
GKW	Guarani Kaiowá
GND	Guarani Ñandeva
GRC	Guarani M'Byá
IP	Índice de Paternidade
KAI	Kaingangs unificados
KIV	Kaingang do Ivaí
KRC	Kaingang do Rio das Cobras
NRC	National Research Council (EUA)
PC	Probabilidade de Coincidência
PD	Poder de Discriminação
PE	Poder de Exclusão
PIC	Conteúdo de Informação Polimórfica
STR	<i>Short Tandem Repeats</i>

## SUMÁRIO

1	INTRODUÇÃO .....	18
2	REVISÃO BIBLIOGRÁFICA .....	21
2.1	MARCADORES STR NA PRÁTICA FORENSE.....	21
2.2	IMPORTÂNCIA DOS ESTUDOS POPULACIONAIS PARA A GENÉTICA FORENSE .....	23
2.3	POPULAÇÕES INDÍGENAS KAINGANG E GUARANI .....	25
2.4	DIVERSIDADE E ESTRUTURA GENÉTICA POPULACIONAL .....	28
2.4.1	Equilíbrio de Hardy Weinberg.....	28
2.4.2	Desequilíbrio de Ligação e Ligação .....	29
2.4.3	Estrutura Populacional.....	30
2.4.3.1	Estatísticas F e R.....	33
2.4.3.2	Fator de Correção ( $\theta$ ).....	37
2.4.3.3	Análise de Variância Molecular (AMOVA) .....	39
2.4.3.4	Structure .....	39
2.4.3.5	Análise de Agrupamento .....	40
2.4.3.6	Análise de Coordenadas (PCoA) e Componentes Principais (PCA) .....	40
2.5	PARÂMETROS ESTATÍSTICOS PARA ANÁLISE FORENSE .....	42
3	OBJETIVOS .....	47
3.1	OBJETIVO GERAL .....	47
3.2	OBJETIVOS ESPECÍFICOS .....	47
4	MÉTODOS .....	49
4.1	AMOSTRAS POPULACIONAIS .....	49
4.2	AMPLIFICAÇÃO E GENOTIPAGEM .....	51
4.4	COMPARAÇÕES COM OUTRAS POPULAÇÕES .....	52
4.3	ANÁLISE DOS DADOS .....	53
5	RESULTADOS E DISCUSSÃO .....	57
5.1	CARACTERIZAÇÃO DA DIVERSIDADE GENÉTICA EM GUARANIS E KAINGANGS: FREQUÊNCIAS ALÉLICAS E EHW .....	57
5.1.2	Desequilíbrio de ligação .....	66
5.2	DISTRIBUIÇÃO DA DIVERSIDADE GENÉTICA MOSTRA ESTRUTURAÇÃO POPULACIONAL ENTRE GUARANIS, MAS NÃO ENTRE KAINGANGS .....	68
5.2.1	AMOVA e Estatística R .....	68
5.2.2	Structure .....	72

5.2.3 Agrupamento hierárquico .....	73
5.2.4 PCoA .....	74
5.4 PARÂMETROS FORENSES .....	77
5.5 DISTRIBUIÇÃO DA DIVERSIDADE GENÉTICA DE GUARANIS E KAINGANGS COM RELAÇÃO A OUTRAS POPULAÇÕES.....	80
5.5.1 Populações Brasileiras Guarani, Kaingang e Urbana Miscigenada.....	80
5.5.1.1 AMOVA e Estatística R.....	98
5.5.1.2 Structure .....	102
5.5.1.3 Agrupamento hierárquico.....	103
5.5.1.4 PCoA.....	104
5.5.2 Populações Guarani e Kaingang, parentais e outros ameríndios.....	106
5.5.2.1 Structure .....	108
5.5.2.2 PCoA e PCA .....	110
5.5.2.3 Agrupamento hierárquico.....	113
5.5.2.4 Estatística F .....	116
5.5.2.5 Reconsiderações sobre os valores de theta para populações ameríndias .....	117
6 CONCLUSÕES .....	121
REFERÊNCIAS .....	123
ANEXO I.....	134
ANEXO II .....	140
ANEXO III .....	141
ANEXO IV .....	142
ANEXO V .....	148
ANEXO VI.....	151

## 1 INTRODUÇÃO

A individualização genética, objetivo da ciência forense, é atualmente realizada através de marcadores hipervariáveis do tipo repetições curtas em série (STR, do inglês, *short tandem repeats*), os quais apresentam diversas vantagens em relação a outros marcadores biológicos usados no passado. Com o objetivo de padronizar as análises comparativas forenses ao redor do mundo, o Departamento Federal de Investigação (FBI do inglês, *Federal Bureau of Investigation*) elaborou o banco de dados denominado sistema combinado de índices de DNA (*CODIS do inglês, Combined DNA Index System*), que utiliza atualmente um número mínimo de 20 marcadores STR de uso na análise forense.

Os exames forenses são realizados através da comparação (confronto) entre os perfis genéticos obtidos para um vestígio, de origem desconhecida, e uma amostra referência, de origem conhecida. Há três possibilidades de resultados a partir de um exame de confronto genético: (1) exclusão, quando não são detectadas coincidências durante o confronto; (2) não exclusão, quando se verifica coincidência total; ou (3) resultado inconclusivo, quando a amostra não produz nenhum perfil viável, desfecho comum para amostras muito degradadas. Quando os resultados geram uma situação de não exclusão, é necessário que o significado da associação encontrada seja avaliado através do cálculo da probabilidade de coincidência (para identificação humana) ou do índice de verossimilhança ou de coincidência (quando não se descarta a chance de amostra questionada e de referência serem originadas da mesma pessoa), a fim de se quantificar o valor da coincidência constatada na análise dos perfis genéticos. Estes cálculos estatísticos que visam a valoração da prova do DNA devem incorporar informações de populações e de bases de dados legítimas que disponibilizam estimativas de frequências alélicas em amostras representativas de diferentes populações, urbanas ou isoladas.

Nesse contexto, o estudo de diferenciação genética envolvendo populações isoladas e de menor tamanho efetivo que as populações urbanas, contribui para se avaliar a adequação dos bancos de dados populacionais

utilizados pelos laboratórios forenses às diretrizes do *National Research Council* (NRC, do inglês, Conselho Nacional de Pesquisa), visando o cálculo do significado estatístico adequado a ser realizado para uma coincidência de DNA.

Com o objetivo de se considerar, para as populações em estudo, a eventualidade de compartilhamento de alelos devido a um ancestral comum, a comunidade forense adota um fator de correção, denominado  $\theta$ , a ser aplicado nos cálculos mencionados. O valor de  $\theta$  reflete a proporção de diversidade genética que se deve a diferenças nas frequências alélicas entre populações, devendo ser integrado aos cálculos de probabilidade de todos os locos. Seu valor é essencialmente o mesmo do índice de coancestralidade  $R_{st}$  (ou  $F_{st}$ ), uma vez que são termos usados de forma equivalente para refletir a estrutura populacional relativa à diversidade genética entre populações ou em relação à população total. Para estudos com STRs forenses, o índice  $R_{st}$  é mais indicado por contemporizar o modelo mutacional apresentado por estes marcadores.

Órgãos reguladores nacionais e internacionais estabeleceram uma padronização dos valores de  $\theta$  como sendo 0,01 para populações urbanas e 0,03 para populações isoladas. No entanto, estudos têm demonstrado que o valor padronizado para populações isoladas pode apresentar variações, com possíveis impactos diretos nas análises forenses, o que sugere que seja mais prudente que se estabeleça o valor deste fator individualmente para cada população.

A diferenciação genética relatada entre diversos grupos populacionais motivou a recomendação de que populações representativas de diferentes ancestralidades deveriam ser utilizadas em estudos populacionais com fins forenses (NRC II, 1996). Embora essas recomendações tenham sido originalmente dirigidas aos Estados Unidos da América, estas também se aplicam ao Brasil, cuja história de miscigenação promoveu uma elevada diversidade genética. Dezenas de bases de dados para frequências alélicas de marcadores STR nas populações urbanas brasileiras já se encontram, desde então, disponibilizadas.

Populações indígenas nas Américas foram descritas como as populações que apresentam o maior nível de subdivisão genética conhecido, o que reflete sensivelmente em suas frequências genotípicas e alélicas. No entanto, bancos de frequências alélicas de marcadores STR específicos entre

indígenas permanecem escassos, no Brasil e no mundo, possivelmente pela dificuldade de acesso a amostras destas populações. Os dados populacionais de que se dispõe atualmente não podem ser considerados representativos da maioria das populações indígenas, o que compromete sua aplicação em rotinas forenses.

Sendo assim, o presente estudo teve como objetivo analisar a diversidade populacional de marcadores de uso forense em populações ameríndias geneticamente isoladas do Brasil e, desta maneira investigar a existência de subdivisão populacional e estimar os valores específicos de  $\theta$  para estas populações. Para tanto foram estimadas as frequências alélicas para 21 marcadores STRs autossômicos em amostras de duas diferentes populações Kaingang e três Guaraní do Paraná e do Mato Grosso do Sul. A avaliação da estruturação incluiu estatísticas R e F, análise de variância molecular, análise de reconstrução e de coordenadas principais. Além destas, foram estabelecidos os parâmetros estatísticos de interesse forense, como a verificação do equilíbrio de Hardy Weinberg, desequilíbrio de ligação, probabilidade de coincidência, poder de discriminação, conteúdo de informação polimórfica, índice de paternidade e poder de exclusão.

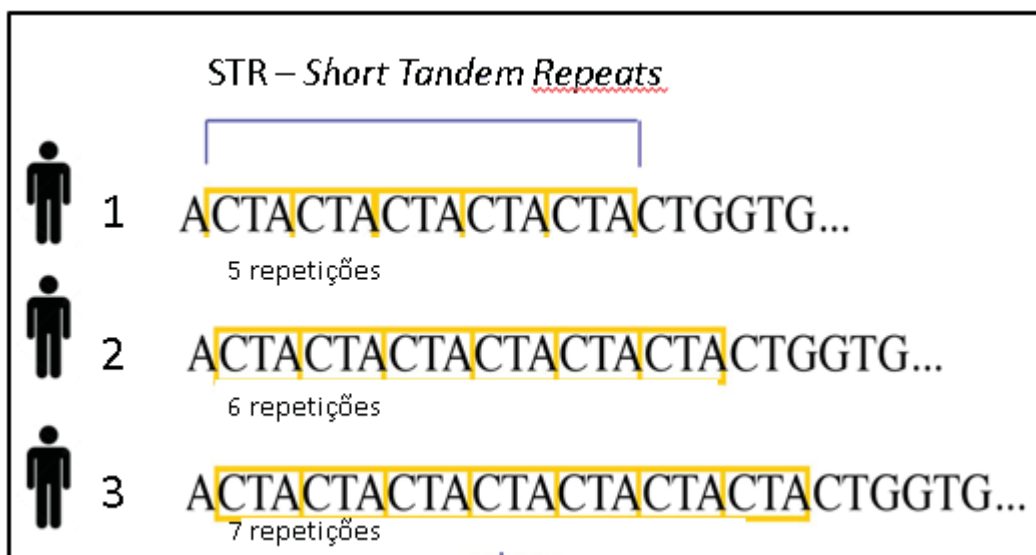


## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 MARCADORES STR NA PRÁTICA FORENSE

A amplificação do DNA por meio da reação em cadeia da polimerase (PCR do inglês, *polymerase chain reaction*), a automação nas técnicas de genotipagem e a descoberta dos STR (do inglês, *short tandem repeats*) possibilitaram a aplicação disseminada de métodos envolvendo a análise de DNA nos laboratórios forenses que visam a identificação humana e o perfil genético de um vestígio originário de um exame de corpo de delito (JOBLING, 2004).

Marcadores STR consistem em um tipo de polimorfismo de comprimento constituído por sequências curtas de nucleotídeos, formando um motivo de dois a sete pares de bases que se repetem consecutivamente, por um número variável de vezes, de forma justaposta e sem apresentar sobreposição. Estes motivos se distinguem pelo seu tamanho e pelo número de vezes que ocorrem numa sequência, para cada marcador. O conjunto de marcadores STR de uso forense oferece a possibilidade de diferenciar indivíduos (Figura 1).



**FIGURA 1** – REPRESENTAÇÃO DE UM MARCADOR STR EM TRÊS INDIVÍDUOS DISTINTOS, CADA QUAL APRESENTANDO DIFERENTES NÚMEROS DE REPETIÇÕES DE TRINUCLEOTÍDEOS.

FONTE: Adaptado de BUTLER, 2012.

Tais sistemas são efetivos para fins de identificação humana inequívoca por apresentarem i) elevado poder de discriminação, ii) tamanho reduzido, o que os torna mais apropriados para amostras degradadas, iii) abundância, já que constituem aproximadamente 3% do genoma humano, iv) onipresença, pois ocorrem em média a cada 10 mil nucleotídeos, v) codominância, vi) universalidade, vii) reprodutibilidade, e viii), por apresentarem um padrão polialélico significativo (devido a inserções, deleções ou mutações) que lhes confere robusta efetividade na análise de misturas de diversas fontes biológicas (JOBLING, 2004; GROVER *et al.*, 2014).

Habitualmente, os marcadores STR empregados atualmente na casuística forense são aqueles constituídos de repetições de tetranucleotídeos (SILVA *et al.*, 2012), por gerarem menor incidência de bandas fantasma, que são artefatos de PCR decorrentes de uma ou duas unidades de repetição, a mais ou a menos, no motivo recorrente presente no alelo original. Tais artefatos podem levar a conclusões errôneas de genotipagem (WALSH *et al.*, 1996).

Em 1997, o Departamento Federal de Investigação (FBI do inglês, *Federal Bureau of Investigation*) selecionou 13 marcadores STR altamente polialélicos e polimórficos para a criação de um banco de perfis genéticos visando a identificação de criminosos, de pessoas desaparecidas, ou de cadáveres. Os marcadores escolhidos passaram a compor o banco de perfis genéticos nacional dos Estados Unidos, denominado Sistema Combinado de Índices de DNA (CODIS, do inglês *Combined DNA Index System*). A padronização destes marcadores possibilitou que todos os laboratórios forenses utilizassem, no mínimo, os mesmos locos, um requisito necessário para a comparação de perfis genéticos (SILVA *et al.*, 2012; BUDOWLE *et al.*, 1999).

Com o intuito de reduzir a incidência de compatibilidades acidentais, proporcionar maior compartilhamento internacional de dados e incrementar o poder de discriminação visando elucidar casos de pessoas desaparecidas, o FBI recomendou, em 2017, a expansão da relação dos marcadores CODIS originais. Os novos locos deveriam atender aos seguintes critérios: (1) não apresentar associação com condições médicas, (2) apresentar baixa taxa de

mutação, (3) alto nível de independência, (4) alto nível de discriminação, (5) possibilitar o uso pela comunidade forense internacional, (6) apresentar um fator de relação de número de locos recomendados com o nível de discriminação oferecido e (7) apresentar conformidade com os requerimentos de qualidade como validação, especificidade humana, dentre outros. Atualmente, o painel de marcadores que compõem o CODIS é constituído pelos locos *CSF1PO*, *FGA*, *TH01*, *TPOX*, *vWA*, *D3S1358*, *D5S818*, *D7S820*, *D8S1179*, *D13S317*, *D16S539*, *D18S51*, *D21S11*, *D2S1338*, *D19S433*, *D1S1656*, *D12S391*, *D2S441*, *D10S1248* e *D22S1045*.

A genotipagem dos marcadores STR, nos laboratórios forenses, costuma ser efetuada através de sistemas comerciais (HARES, 2012). Dentre os sistemas comerciais disponíveis, o sistema *GlobalFiler* (*Thermo Fisher Scientific, Waltham, Massachusetts, EUA*), oferece a amplificação simultânea dos 20 marcadores CODIS, além do loco SE33 e de mais três marcadores para identificação de sexo do fornecedor da amostra (*Amelogenina* e os marcadores *DYS391* e *Yindel*). Assim, este sistema vem sendo rotineiramente utilizado nos laboratórios forenses.

## 2.2 IMPORTÂNCIA DOS ESTUDOS POPULACIONAIS PARA A GENÉTICA FORENSE

A genética molecular forense baseia-se, invariavelmente, no confronto entre os perfis genéticos da evidência (de origem questionada) e de uma amostra referência (de origem conhecida). Denomina-se aqui amostra evidência (ou vestígio) toda amostra oriunda de um exame de corpo de delito. Do ponto de vista pericial, entende-se como corpo de delito a materialidade de um crime. Isto é, toda expressão do delito em sua corporação física, como o corpo humano, o local de um crime e seus vestígios.

O confronto genético pode gerar três resultados possíveis: (1) exclusão, quando são detectadas ausências de coincidências entre os alelos em parte ou em todas as regiões analisadas durante o confronto; (2) não exclusão, quando se verifica coincidência total; ou (3) resultado inconclusivo, quando a amostra não produz nenhum perfil viável, como é comum para amostras muito degradadas (COLLINS, 1994). Quando um exame de confronto produzir

resultados de não exclusão, e apenas nesta situação, será necessário que se quantifique o valor da associação constatada na análise dos perfis genéticos, através da probabilidade de coincidência (para a identificação humana) e do índice de verossimilhança (quando não se descarta a chance de amostra questionada e de referência serem originadas da mesma pessoa), brevemente explicados a seguir.

A probabilidade de que o perfil genético obtido da amostra questionada seja encontrado em um indivíduo aleatoriamente escolhido na população é denominado probabilidade de coincidência. Este valor é obtido a partir das frequências do perfil genético nas populações e responde basicamente ao questionamento de qual seria a probabilidade de uma pessoa aleatoriamente escolhida na população apresentar este mesmo perfil genético.

O índice de verossimilhança consiste na medida da força da evidência considerando duas hipóteses exaustivas e mutuamente excludentes, habitualmente denominadas hipótese da acusação e hipótese da defesa. Na hipótese da acusação, supõe-se que o perfil genético encontrado na amostra evidência seja oriundo do suspeito. Na hipótese da defesa, parte-se do pressuposto de que o perfil genético encontrado na amostra evidência seria idêntico ao perfil do suspeito por uma mera coincidência e que, na realidade, seria proveniente de um indivíduo aleatório e não testado da população (CURRAN *et al.*, 2007).

Estes cálculos estatísticos que têm como objetivo a valoração da prova do DNA deverão incorporar dados de frequências alélicas obtidas a partir da contagem de alelos em amostras representativas de populações. Com o objetivo de conferir a maior precisão possível a estes cálculos, é imperativo que se considere a diversidade genética (frequência dos alelos STR) na população, inclusive de grupos populacionais minoritários, uma vez que as frequências alélicas dos marcadores forenses podem variar nas diversas populações (PRITCHARD *et al.*, 2000).

Através das estimativas populacionais de frequência alélica é possível se estimar o grau de raridade com o qual um alelo pode ser encontrado, além de obter informações sobre a existência de alelos exclusivos a uma determinada população ou, ainda, quantificar a diversificação genética existente entre duas populações (CURRAN *et al.*, 2007; TAMURA *et al.*, 2015).

Estimativas de frequências alélicas têm sido conduzidas ao redor do mundo, abrangendo populações urbanas e, em menor proporção, populações isoladas. Estudos de diferenciação genética envolvendo populações isoladas e de menor tamanho efetivo que as populações urbanas contribuem para a adequação dos cálculos realizados pelos laboratórios forenses ao redor do mundo, às diretrizes do Conselho Nacional de Pesquisa (NRC, do inglês, *National Research Council*). Tais estimativas devem ser criteriosas considerando-se seu impacto nos cálculos estatísticos envolvidos em investigações de vínculo genético. Foi descrito que populações ameríndias apresentam o maior nível de subdivisão populacional conhecido, o que diferencia sensivelmente suas frequências genotípicas e alélicas com relação a outras populações indígenas (WANG *et al*, 2007; BUDOWLE *et al*, 2001, McCULLOH *et al*, 2016). O NRC recomenda que, na ausência de dados específicos a uma população indígena, se utilize dados de uma população indígena que se julgue ser similar à primeira (NRC II, 1996). Contudo, a diferenciação existente para uma dada população, caso não seja considerada, pode comprometer o cálculo do significado estatístico adequado a ser realizado para uma coincidência de DNA que envolva estas populações (MANTA *et al.*, 2011). No entanto, bancos de frequências alélicas de marcadores STR específicos entre indígenas permanecem escassos, no Brasil e no mundo, possivelmente em virtude da dificuldade de acesso a amostras destas populações (KANTHASWAMY *et al*, 2014). Dessa maneira, os dados populacionais de que se dispõe atualmente não podem ser considerados representativos da maioria das populações ameríndias e seria mais desejável que não fossem considerados de forma generalizada nos cálculos envolvidos em rotinas forenses.

### 2.3. POPULAÇÕES INDÍGENAS KAINGANG E GUARANI

Estima-se que, dos mais de 1.000 povos nativos americanos que habitavam o território brasileiro antes da povoamento europeia, apenas cerca de 200 continuam existindo, constituídos por pouco mais de 817 mil indivíduos remanescentes (censo demográfico IBGE – Instituto Brasileiro de Geografia e

Estatística – 2010). A maior parte dessa população encontra-se distribuída por aldeias situadas no interior de reservas indígenas espalhadas do norte ao sul do território nacional (Instituto Socioambiental, 2016), sendo subdivididas de acordo com aspectos de suas culturas (MARRERO, 2007).

Nove povos indígenas são atualmente encontrados no Paraná, distribuídos em 26 terras indígenas, onde habitam populações Kaingang, Guarani (Ñandeva e M'byá) e Xetá. No Mato Grosso do Sul, existem 53 terras indígenas, nas quais residem Guarani (Kaiowá e Ñandeva) (INSTITUTO SOCIOAMBIENTAL). Como será apresentado a seguir, são grandes as barreiras culturais entre os povos Guarani e Kaingang, o que manteve essas populações geneticamente isoladas ao longo do tempo (PETZL-ERLER, *et al.*, 1993).

Na época do primeiro contato com os europeus os indígenas Guarani habitavam um vasto território na porção sul do Brasil, incluindo São Paulo, Rio de Janeiro, Mato Grosso do Sul, Paraná e Rio Grande do Sul, além da porção oeste do Paraguai, Bolívia, nordeste e centro oeste da Argentina (região das províncias de Misiones, Corrientes e Entre Rios) (SALA *et al.*, 2010; TSUNETO *et al.*, 2003). Os Guarani eram nômades, procuravam planícies próximas de grandes rios e pertenciam ao grupo linguístico Tupi-Guarani, que faz parte do ramo Tupi. Os primeiros contatos com europeus (espanhóis e portugueses) datam do século XVI após o qual parte da população foi dizimada enquanto o restante foi incorporado à população miscigenada brasileira ou, em parte, permaneceu isolado (TSUNETO *et al.*, 2003). Hoje os Guarani representam a maior etnia indígena brasileira, sendo divididos em três grupos de acordo com o dialeto que falam e diferindo em diversos aspectos de sua cultura (Guarani M'byá, Guarani Ñandeva e Guarani Kaiowá). Tais diferenças cumprem um papel de marcadores étnicos e segregam os grupos tanto entre si quanto com relação a outras etnias (indígenas ou urbanas), mesmo que vivam em áreas compartilhadas com outras comunidades indígenas ou próximas a centros urbanos. Dentre os diferentes grupos, o mais tradicional na preservação de sua identidade cultural é o Guarani M'Byá. Atualmente, os Guarani encontram-se distribuídos em reservas indígenas no sul e sudeste do Brasil, no Paraguai, norte da Argentina e na Bolívia, frequentemente nas proximidades de povos Kaingang. A principal atividade econômica é a agricultura, seguida de caça,



pesca e artesanato. Permanecem em comunidades bastante restritas, de modo a preservar suas organizações sociopolíticas e econômicas (INSTITUTO-SOCIOAMBIENTAL, 2016; PETZL-ERLER, *et al.*, 1993).

O termo Kaingang (anteriormente Guayaná e Coroadó, dentre outros nomes) foi introduzido em 1882 para designar todos os povos não Guaraní que viviam no sul do Brasil (MARRERO *et al.*, 2007). Os Kaingang pertencem ao grupo linguístico Jê, do tronco Macro-Jê. Escolhiam planaltos para habitar e apresentavam hábitos semi-nômades. Praticavam a exogamia patrilinear que restringia incesto, sendo proibido o casamento entre indivíduos da mesma subtribo. A poligamia era frequente. Os contatos com europeus datam de antes da metade do século XVIII. Além de um histórico permeado por violência com os parentais brancos e também entre si, há também relatos posteriores de que muitas destas populações indígenas teriam sido atingidas por epidemias. Esta história pregressa gerou significativo impacto demográfico. Vivem atualmente em 32 territórios indígenas e caracterizam atualmente a terceira maior etnia indígena do Brasil, com uma população estimada (os censos são precários porque as famílias Kaingang mudam-se frequentemente de aldeia) de aproximadamente 25.000 indivíduos. Suas principais atividades econômicas são agricultura, caça, coleta de frutos e artesanato (INSTITUTO-SOCIOAMBIENTAL, 2016; PETZL-ERLER, *et al.*, 1993).

Os povos Guaraní e Kaingang vivem em vilas separadas, embora compartilhando por séculos os territórios destinados aos indígenas. Essa proximidade geográfica prolongada não foi capaz de romper a significativa barreira imposta pela língua falada por cada grupo indígena. A linguagem representa uma barreira cultural de expressiva magnitude, exercendo um obstáculo ao fluxo gênico tão ou mais significativo que a distância geográfica (FAGUNDES *et al.*, 2002; KOHLRAUSCH *et al.*, 2005). Guaranis e Kaingangs permanecem diferenciando-se claramente entre si e cultivando hábitos predominantemente endogâmicos que trazem, como reflexo, o respectivo isolamento populacional (PETZL-ERLER *et al.*, 1993; BELTRAME *et al.*, 2008; MARRERO *et al.*, 2007).

Até o momento, foram estudados aqueles marcadores STRs, inclusive STRs forenses integrantes do painel original no CODIS, em determinadas populações indígenas brasileiras, incluindo algumas populações Guaraní e

Kaingang. Tais estudos, contudo, consideraram as populações Guarani, de localidades diferentes, como uma única população Guarani. Da mesma forma, todas as populações Kaingang oriundas de regiões geográficas diversas foram analisadas como sendo uma população Kaingang única (KOHLRAUSCH *et al*, 2005). No presente trabalho foram analisadas três populações Guarani e duas Kaingang dos estados do Paraná e Mato Grosso do Sul, que representam duas das principais populações indígenas do Brasil. As respectivas populações, provenientes de regiões geograficamente distintas, foram tratadas e analisadas individualmente.

## 2.4 DIVERSIDADE E ESTRUTURA GENÉTICA POPULACIONAL

A estimativa das frequências genotípicas e alélicas populacionais serve de base para as medidas de diversidade genética, estrutura populacional e parâmetros de interesse forense. As frequências dos perfis genéticos são estimadas a partir de estudos populacionais, sendo então disponibilizadas através de bancos de frequências alélicas. Para que sejam considerados eficazes para fins forenses, os parâmetros investigados são equilíbrio de Hardy-Weinberg, desequilíbrio de ligação entre os marcadores STR analisados, análise de variância molecular (AMOVA), distância e estrutura populacional (estatística R, Structure, análise de componentes principais, agrupamento hierárquico), heterozigosidade observada e esperada, poder de discriminação (PD), conteúdo de informação polimórfica (PIC), probabilidade de coincidência (PC), poder de exclusão (PE), índice de paternidade (IP), os quais serão explicados e suas importâncias discutidas a seguir.

### 2.4.1 *Equilíbrio de Hardy Weinberg*

Um dos princípios mais importantes para os estudos de variação genética foi proposto por Hardy e Weinberg em 1908, sendo conhecido como princípio ou equilíbrio de Hardy-Weinberg. Por este princípio, assumindo-se pressupostos como população de tamanho infinito, cruzamentos ao acaso e ausência de fatores evolutivos que alteram as frequências alélicas (seleção, deriva genética, fluxo gênico e mutação), as frequências genotípicas e alélicas permanecerão constantes ao longo do tempo. À essa constância de

frequências denomina-se de equilíbrio de Hardy-Weinberg (EHW). Qualquer violação dos pressupostos citados pode produzir resultados que afetam as frequências alélicas e genotípicas. O EHW consiste em uma premissa teórica utilizada como modelo referencial no qual não existiria a atuação de forças evolutivas. Em populações reais e, especificamente para a maioria dos STRs de interesse forense, a atuação de fatores evolutivos é baixa.

As frequências genotípicas podem ser inferidas a partir das frequências alélicas (RYCKMAN, 2008; TRACEY, 2001). As frequências genotípicas esperadas (de acordo com estimativas) e observadas são comparadas e, se estiverem próximas, assume-se que a população está em EHW. Nesta situação, as combinações alélicas são consideradas independentes umas das outras.

O valor da prova de DNA na genética forense reside na certeza de independência estatística nos cálculos realizados em caso de coincidência de perfis. Logo, na casuística forense, para que um loco seja considerado plenamente apropriado para investigações genéticas de identificação humana ou de probabilidade de parentesco, é importante que se estabeleça se este se encontra em EHW (TRACEY, 2001; LANGE, 1993). Os desvios das frequências genotípicas esperadas no EHW podem sugerir a ocorrência de endogamia e estruturação populacional, ou podem representar um fator de ordem técnica (e não evolutiva) como os decorrentes de erros de amostragem e ou do método utilizado para fazer a genotipagem, incluindo falhas na amplificação, como *drop-out*, ou amplificação preferencial, como *drop-in* (WITTKE-THOMPSON *et al.* 2005; SCHNEIDER, 2007).

#### 2.4.2 Desequilíbrio de Ligação e Ligação

Parte da estatística utilizada para o cálculo de probabilidade de coincidência se baseia na *regra dos produtos*. Isto significa que as probabilidades de coincidência calculadas individualmente para cada marcador STR devem ser multiplicadas entre si visando reduzir a chance de uma coincidência aleatória. Para que esse cálculo seja empregado de forma estatisticamente correta e coerente, é necessário que os marcadores envolvidos apresentem herança independente uns dos outros.

Tradicionalmente, as investigações na genética molecular forense fundamentam-se em marcadores genéticos que se encontrem em equilíbrio de ligação, isto é, alelos que sejam herdados de forma independente dos alelos de outros locos. Quando se verifica a herança independente, as frequências alélicas observadas em uma população se mostram próximas às frequências esperadas e a regra dos produtos pode ser usada nos cálculos forenses, como recomendado (NRC II, 1996).

Por desequilíbrio de ligação (DL) entende-se a associação não aleatória de diferentes marcadores. Dentre as possíveis causas do DL estão a proximidade física (ligação) entre genes sintênicos e fatores demográficos (TAMURA *et al.*, 2015; TILLMAR *et al.*, 2015).

Algo importante a ser esclarecido é que os termos “ligação” e “desequilíbrio de ligação” são frequentemente confundidos e muitas vezes considerados erroneamente como sinônimos. Esse equívoco acontece porque a ligação pode ser uma das causas do DL. Entretanto, o fato de dois locos estarem em DL não significa que estejam ligados. Por esta razão, muitos autores recomendam o uso do termo desequilíbrio gamético para DL (GUPTA *et al.*, 2005).

Denomina-se “ligação” (do inglês, *linkage*) a ocorrência de segregação conjunta de marcadores fisicamente localizados em regiões próximas, em um mesmo cromossomo. A possibilidade da ocorrência de recombinação entre eles depende da distância que os separa. Quanto mais próximos, maior a possibilidade de que sejam herdados em bloco, gerando associação não aleatória de alelos.

Fatores da dinâmica populacional, como a subestruturação populacional e em particular a deriva genética (como o efeito fundador), ou ainda fatores como ancestralidade, podem influenciar na ocorrência de DL, requerendo que este seja investigado em populações isoladas como as deste estudo (LEWONTIN, 1988, GUPTA *et al.*, 2005; PAYSEUR *et al.*, 2008 e NORDBORG *et al.*, 2008).

#### 2.4.3 Estrutura Populacional

Estrutura populacional consiste na presença de níveis hierárquicos de subpopulações que levam à diversidade genética, seja por processos evolutivos (como seleção natural, deriva, migração) ou aleatórios (como mutação, localização geográfica, diversificação social). Em termos práticos, a estrutura populacional reflete o número de alelos trocados entre subpopulações, influenciando a composição genética dos indivíduos que as compõem. Processos como seleção e deriva genética aumentam a diversidade genética entre as populações em graus variáveis. Paralelamente, os processos demográficos opostos, como a migração e a miscigenação, reduzem a diversidade genética entre populações, homogeneizando as frequências dos alelos entre elas. Populações, sobretudo aquelas de tamanho reduzido, podem ainda estar sujeitas ao endocruzamento (transmissão de alelos idênticos por descendência) e, conseqüentemente à redução da diversidade genética (ELHAIK, 2012; BALLOUX *et al*, 2002; TVEDEBRINK, 2010; CHAKRABORTY, 1992; MCCULLOH *et al.*, 2016; MANTA *et al.*, 2011).

Todas as populações costumam estar sujeitas a algum nível de estruturação genética. A formação da população brasileira, em particular, se deve à misturas decorrentes de ondas migratórias começando com a população ameríndia, prosseguindo pela ocupação pelos parentais portugueses, os anos de imigração de africanos como mão de obra escrava, além de imigração posterior da Itália, Alemanha, Espanha e Japão, dentre outras, originando uma população altamente miscigenada. O resultado de todo este processo demográfico, através de ondas migratórias de distribuição desigual ao longo do país, foi o estabelecimento de um padrão heterogêneo de miscigenação entre europeus, africanos e ameríndios (LEITE *et al*, 2009; SCHWENGBER, *et al*, 2009). Desta forma, as populações urbanas brasileiras apresentam componentes genéticos oriundos de diversas ancestralidades e subpopulações.

O Brasil é um país com proporções continentais, mas esta configuração populacional pode ser facilmente constatada em áreas de menor abrangência, como as cidades, nas quais é comum encontrar bairros com predomínio de diferentes ancestralidades.

Na genética forense, os cálculos de probabilidade consideram as variações na constituição genética de indivíduos na população, requerendo,

portanto, o conhecimento a respeito da respectiva estrutura populacional (CURRAN *et al.*, 1999, TVEDEBRINK, 2010; ELHAIK, 2012). No Brasil, onde a estruturação populacional não é homogênea, torna-se necessário que esta seja identificada e dimensionada para diversas populações.

Até o momento a comunidade científica já reuniu dados sobre as frequências dos marcadores de interesse forense obtidos para diversas populações no mundo, bem como para diferentes populações brasileiras. A maior parte destes dados refere-se a populações urbanas. No Brasil, tais dados foram constituídos por amostragens representativas da miscigenação característica do país (GRATAPAGLIA *et al.*, 2001; AGUIAR *et al.*, 2012; FRANCEZ *et al.*, 2011; OLIVEIRA *et al.*, 2014; RODRIGUES *et al.*, 2009; SILVA *et al.*, 2004; MANTA *et al.*, 2011; GUSMÃO *et al.*, 2016).

Populações presentes nas Américas, inclusive populações indígenas brasileiras, apresentam, de maneira geral, significativa escassez de informações específicas sobre diversidade genética (MCCULLOH *et al.*, 2016; MANTA *et al.*, 2011). Consequentemente, são escassos os dados populacionais de ameríndios para aplicação forense, possivelmente em virtude da dificuldade em se obter amostras representativas destas populações (KANTHASWAMY *et al.*, 2014). Considerando que a diferenciação genética entre populações possui um componente geográfico e cultural importante, estudos de diversidade de populações que sejam geográfica e culturalmente representativas são de grande interesse e aplicabilidade na comunidade forense (MCCULLOH *et al.*, 2016). É o caso de populações dentro das denominações Guarani e Kaingang, que apresentam distinções culturais e geográficas entre si e com relação à população brasileira miscigenada.

Com o objetivo de facilitar as análises de informações geradas por estudos populacionais, tais como as frequências genotípicas e alélicas, ou ainda dados geográficos, linguísticos, culturais ou físicos, estas podem ser comparadas e agrupadas. Uma metodologia para tal consiste em agrupá-las em classes nas quais possam ser consideradas semelhantes entre si e diferentes de elementos de outra (s) classe (s). O conhecimento prévio das classes não é requisito necessário, visto que programas estatísticos possuem a habilidade de gerá-las por métodos dinâmicos ao longo do processo classificatório (TENDEIRO, 2005; ZARE CHAHOUKI, 2011).



As comparações dos dados populacionais com objetivo de investigar a existência de estrutura populacional podem ser efetuadas através de metodologias estatísticas como, por exemplo, as estatísticas F e R. Análises através de agrupamentos, e outras análises análogas, podem ser realizadas utilizando-se, por exemplo, dendrogramas e análises de componentes e coordenadas principais. Estes métodos, efetivamente utilizados para a realização das análises efetuadas neste trabalho, encontram-se, brevemente, explanados a seguir.

Na falta de processos sistemáticos disponíveis para a escolha do método de análise que melhor se aplica a cada estudo, entende-se que a melhor conduta é integrar as diversas metodologias disponíveis e, ciente das limitações de cada uma, associando os resultados que tenham sido obtidos a partir de um mesmo conjunto de dados (TENDERO, 2005).

#### 2.4.3.1 Estatísticas F e R

A diferenciação genética observada entre populações é passível de ser mensurada através da estatística F de Wright, que fornece um índice de fixação (F) a fim de estimar a conectividade, padrões de fluxo gênico, quantificar o efeito do endocruzamento da subdivisão populacional e, assim, comparar de forma objetiva o efeito das estruturas populacionais (WRIGHT, 1965). A estatística F é calculada a partir da heterozigosidade média das populações, sendo constituída por três parâmetros que avaliam correlações entre (1) a proporção de endogamia para um indivíduo relativamente à população total ( $F_{IT}$ ), (2) a proporção de endogamia para um indivíduo relativamente à subpopulação à qual ele pertence ( $F_{IS}$ ) ou, ainda, (3) a diferenciação de uma subpopulação com relação a uma população total ( $F_{ST}$ ) (MCCULLOH *et al.*, 2016; HOLSINGER *et al.*, 2009; WRIGHT, 1965; BALLOUX *et al.*, 2002).

Na investigação das forças evolutivas que possam atuar na diminuição da heterozigosidade, como endogamia, é indicada a utilização do coeficiente de endogamia ( $F_{IS}$ ). Basicamente, o parâmetro  $F_{IS}$  quantifica o quanto da probabilidade de dois alelos idênticos por descendência se deve, exclusivamente, aos efeitos de endogamia.  $F_{IS}$  é obtido a partir das médias

das frequências esperadas e observadas de heterozigosidade das amostras populacionais em estudo. Quando seu valor é igual a zero, conclui-se pela ausência de endogamia e as frequências gênicas e genotípicas estariam de acordo com as esperadas na condição de equilíbrio de Hardy-Weinberg. Quando seu valor é igual a um, subentende-se que a endogamia é absoluta, com presença exclusiva de genótipos homozigóticos para cada um de seus alelos:

- $F_{IS} = 0$  indica ausência de processos endogâmicos atuantes;
- $F_{IS} > 0$  indica excesso de homozigotos em relação ao esperado pelo modelo de EHW;
- $F_{IS} < 0$  indica excesso de heterozigotos em relação ao esperado pelo EHW.

O parâmetro mais usado na casuística forense,  $F_{ST}$ , é a medida mais inclusiva de subdivisão populacional e fornece os subsídios para se examinar o nível total de divergência genética (variação na frequência de alelos) entre subpopulações com relação à população total.  $F_{ST}$  equivale à redução da heterozigosidade esperada sob acasalamientos ao acaso, comparativamente, em uma subpopulação, e outro nível, qualquer, mais inclusivo, como a população como um todo. Através de sua determinação, portanto, é possível que se quantifique as diferenças genéticas entre subpopulações (ELHAIK, 2012).

Os valores de  $F_{ST}$  variam de 0 a 1, sendo que o valor igual a zero indica a não existência de diferenciação genética, enquanto que o valor de 1 indica a fixação de alelos alternativos em diferentes subpopulações. Normalmente estima-se que os valores de  $F_{ST}$  traduzem diretamente o grau da diferenciação, sendo considerados os seguintes intervalos de diferenciação quando considerados marcadores STR (BALLOUX *et al*, 2002):

- 0-0,05: diferenciação genética pequena.
- 0,05-0,15: diferenciação genética moderada.
- 0,15-0,25: diferenciação genética grande.
- $>0,25$ : diferenciação genética muito grande.

A média e a variância de  $F_{ST}$  dependem de vários fatores, como frequências de alelos, o tamanho e subdivisões das populações e processos

demográficos, e são difíceis de se prever analiticamente na ausência da distribuição de  $F_{ST}$  do genoma total. Como resultado, o  $F_{ST}$  médio, calculado a partir de um subconjunto da distribuição de  $F_{ST}$ , como por exemplo aquele obtido para marcadores STR forense, é frequentemente utilizado para quantificar a divergência genética global entre populações humanas (ELHAIK, 2012). Eventos como deriva, seleção e endogamia, nas populações em estudo, podem aumentar o valor de  $F_{ST}$  enquanto que fluxo gênico tende a diminuí-lo.

A aplicação de  $F_{ST}$ , contudo, pode apresentar limitações quando se trata de STRs, devido às taxas mutacionais envolvidas (HOLSINGER *et al.*, 2015; GAGGIOTTI *et al.*, 1999). Polimorfismos de tamanho, em STRs, devem-se principalmente a mecanismos de deslizamento da polimerase, enquanto que o tamanho de novos alelos depende do tamanho do alelo submetido à mutação (GAGGIOTTI *et al.*, 1999; FAN *et al.*, 2007; LECLERCQ *et al.*, 2010).

Deslizamentos da polimerase ocorrem durante a replicação do DNA, ocasionando um mal pareamento (de uma ou mais repetições) entre a fita molde e a fita nascente, com a formação de um laço neste local, que pode sofrer excisão através de mecanismos de reparo. Como consequência, o número de repetições na fita nascente será alterado para mais ou para menos unidades de repetição que o número original. A taxa de mutação dos STRs depende da taxa de deslizamentos da polimerase e da eficiência do sistema de reparo. A taxa de deslizamentos é mais frequente em STRs com tamanhos menores da unidade de repetição e em STRs mais longos (FAN *et al.*, 2007; LECLERCQ *et al.*, 2010).

Embora não haja um consenso a respeito de um modelo específico que possa descrever com acurácia a dinâmica mutacional em STRs, foram propostos dois modelos para explicar tais eventos mutacionais: (1) o “modelo de alelos infinitos” (IAM do inglês, *infinite alleles model*) e (2) o “modelo gradativo” (SMM do inglês, *stepwise mutation model*) (BALLOUX *et al.*, 2002; GAGGIOTTI *et al.*, 1999), brevemente descritos a seguir:

No IAM entende-se que cada nova mutação produziria um novo alelo, todas as mutações teriam probabilidade equivalente e alelos idênticos compartilhariam a mesma ancestralidade sendo, portanto, idênticos por descendência (homoplasia). Uma variante deste modelo, chamado de “modelo de alelos-K” (KAM, *K-allele model*), propõe que os alelos seriam idênticos por

estado e não por descendência (BALLOUX *et al*, 2002; GAGGIOTTI *et al*, 1999; FAN *et al*, 2007).

No SMM, cada mutação criaria, com igual probabilidade, um novo alelo, seja pela adição ou deleção de uma única unidade de repetição do microssatélite. O “modelo de duas fases” (TPM, *two phases model*), um desmembramento do SMM, foi desenvolvido de forma a considerar a proporção da adição ou deleção de uma ou várias unidades de repetição. Como consequência, alelos de tamanhos diferentes teriam uma relação mais distante entre si do que alelos com tamanhos similares. Por essa razão, subentende-se que SMM considera o tamanho do alelo, diferentemente de IAM e KAM. Para SMM ou TPM, a taxa e tamanho das mutações não apresenta dependência para com a unidade de repetição em si. Nestes modelos (SMM e TPM) há concordância com o mecanismo de deslizamento da polimerase, atualmente o mais aceito para explicar as mutações em STRs (BALLOUX *et al*, 2002; GAGGIOTTI *et al*, 1999).

O uso de  $F_{ST}$  se aplica bem a marcadores sujeitos ao IAM, ou quando a probabilidade de mutação puder ser considerada desprezível. A maioria dos STRs, todavia, está sujeita ao SMM. Para contornar este viés, Slatkin (1995) propôs um outro parâmetro, denominado  $R_{ST}$ , que se baseia na variação do número de repetições e no padrão de mutação (SMM) mais observado nos STRs (BALLOUX *et al*, 2002; GAGGIOTTI *et al*, 1999). Basicamente, o parâmetro  $F_{ST}$  baseia-se na variância das frequências alélicas enquanto que  $R_{ST}$  prioriza a variação no tamanho dos alelos levando em conta as diferenças moleculares entre eles (WANG *et al*, 2001; SILVA *et al*, 2012).

Embora  $F_{ST}$  e  $R_{ST}$  sejam os métodos estatísticos mais utilizados para se estimar estrutura populacional, ambos apresentam, contudo, vantagens e desvantagens. Estimativas baseadas em  $F_{ST}$  não são influenciadas pela diferença amostral. Por outro lado, podem perder precisão quando o número de alelos é excessivamente grande ou pequeno (WANG *et al*, 2001; SILVA *et al*, 2009).

A desvantagem de  $R_{ST}$  ocorre no caso da distribuição não normal de alelos, situação na qual as diferenças entre populações díspares que se aproximam do mesmo tamanho médio podem ser minimizadas (WANG *et al*, 2001). Ainda assim, contudo, o uso de  $R_{ST}$  tem sido considerado um

parâmetro mais apropriado para a análise de STR, assim como os demais parâmetros similares aos da estatística F, como  $R_{IS}$  e  $R_{IT}$ . Seu uso permanece indicado principalmente para populações acima de cinquenta indivíduos e acima de dez marcadores. Quando se trata de amostras de até dez indivíduos, contudo, e até cinco marcadores, as estimativas baseadas em  $F_{ST}$  ainda têm sido melhores do que as baseadas em  $R_{ST}$  (GAGGIOTTI, *et al.*, 1999; BALLOUX, *et al.*, 2002).

#### 2.4.3.2 Fator de Correção ( $\theta$ )

Na rotina forense, quando é realizada a comparação de frequências alélicas de amostras evidência e referência, é importante que a força da evidência pericial seja quantificada. Para tanto, determina-se a probabilidade de coincidência aleatória, isto é, a probabilidade de que o perfil genético obtido a partir de uma amostra de local de crime corresponda ao suspeito no caso do perfil não ser proveniente do suspeito. Esse cálculo é efetuado considerando-se as frequências alélicas da população. É possível, contudo, que o suspeito e o autor do crime pertençam a uma subpopulação para a qual não exista informação de frequências alélicas. Evidências coletadas de um exame de corpo de delito em um crime sem autoria definida se encaixam nesta possibilidade. Em tais situações deve-se utilizar um fator de correção,  $\theta$ , para calcular a probabilidade de uma coincidência, baseada em informações de frequências alélicas de uma população maior da qual a subpopulação faça parte. A probabilidade de coincidência aleatória leva em consideração a variação de frequências alélicas entre subpopulações dentro da população maior cujas frequências alélicas estejam disponíveis. Os valores de  $\theta$  devem ser incorporados aos cálculos de cada loco individualmente.

Basicamente,  $\theta$  reflete a proporção de diversidade que se deve a diferenças nas frequências alélicas entre subpopulações (HOLSINGER *et al.*, 2009; GILL *et al.*, 2014; TRACEY, 2001). O valor de  $\theta$  é essencialmente o mesmo que  $R_{ST}$  (ou  $F_{ST}$ ), sendo ambos os termos usados de forma equivalente para refletir a estrutura populacional relativa à diversidade genética entre populações ou em relação à população total (HOLSINGER *et al.*, 2009; GILL *et al.*, 2014; TRACEY, 2001). O valor de  $R_{ST}$  é mais similar ao valor de  $\theta$  definido

por Weir & Cockerham (1984) do que pelos definidos por heterozigotes pelo fato de ambos serem baseados em componentes de variância interpoblacional (SLATKIN, 1995). Basicamente, se o valor de  $\theta$  ou  $R_{ST}$  (ou  $F_{ST}$ ) for baixo (próximo de zero), compreende-se que a frequência alélica dentro das populações em análise é semelhante e, se o valor for alto, é um indicativo de que as frequências alélicas são diferentes entre as populações, ou seja, que haja alta diferenciação interpoblacional (HOLSINGER *et al.*, 2009).

O fator de correção  $\theta$  deve sempre ser utilizado em dados populacionais quando a existência de subgrupos de populações (ou populações sujeitas a endogamia) não puder ser descartada. As dificuldades em se definir subpopulações relevantes a casos específicos e a falta de amostras representativas destas populações dificultam o estabelecimento de valores individualizados para coeficientes como  $\theta$ . Populações reais nunca são completamente separadas pois sempre haverá a possibilidade de níveis de estrutura geográfica ou social (CURRAN *et al.*, 2007; SCHNEIDER, 2007). Com o objetivo de incorporar de forma conservadora o valor de  $\theta$  aos cálculos envolvidos na casuística forense, o NRC recomenda um valor de  $\theta$  de 0,01 a ser utilizado no caso de populações urbanas dos Estados Unidos. Já para populações menores e sujeitas a maior isolamento, como é o caso dos nativos norte-americanos, situação na qual a subdivisão é mais prevalente, o valor sugerido para  $\theta$  é de 0,03 (NRC II, 1996). A mesma recomendação vem sendo utilizada, pelos laboratórios de genética, para estudos de casos forenses envolvendo a população brasileira.

Na casuística forense é vantajoso para o réu que se use valores elevados para  $\theta$  para reduzir o peso aparente da evidência genética (CURRAN *et al.*, 2007). O efeito prático da correção de  $\theta$  é que a força numérica da evidência contra um suspeito é reduzida. Portanto, é muito menos surpreendente ver uma correspondência quando levamos em conta a subestrutura da população do que quando a ignoramos. Por isso, existe um questionamento se o valor de  $\theta$  poderia estar sendo superestimado, tendo sido constatadas discrepâncias entre os valores de  $\theta$  obtidos por diferentes autores para as mesmas populações (WANG *et al.*, 2007, MacCULLOH *et al.*, 2016). Assim, obter os valores mais acurados para o fator de correção nas diversas

subpopulações é de grande interesse para que à prova pericial seja atribuído o valor mais justo e correto.

A ausência de dados consistentes abre espaço para especulações tendenciosas, de forma que uma abordagem cientificamente validada para estimar estes parâmetros se faz necessária. Assim, dados disponibilizados na literatura devem ser avaliados com rigor científico para fins de determinação de  $\theta$  (CURRAN *et al.*, 2007; TVEDEBRINK, 2010).

#### 2.4.3.3 Análise de Variância Molecular (AMOVA)

A análise de variância molecular (AMOVA) (EXCOFFIER *et al.*, 2005) consiste em um método que, através de valores de  $F_{ST}$  e ou  $R_{ST}$ , permite análises hierárquicas de três componentes da variância genética: (1) diferenças genéticas entre indivíduos dentro das populações ( $V_a$ ); (2) diferenças genéticas entre populações dentro dos grupos ( $V_b$ ); e (3) diferenças genéticas dentro das populações ( $V_c$ ). Esta classe de medidas permite uma estimativa da distribuição dos alelos (STRs, no presente caso) dentro e entre populações, bem como a distribuição de distâncias evolucionárias entre os alelos (HOLSINGER *et al.*, 2009).

#### 2.4.3.4 Structure

O programa STRUCTURE é normalmente utilizado para inferir o número de grupos ( $K$ ) em que um conjunto de populações pode ser organizado. O número de grupos formados pode ser predeterminado, porém são os dados fornecidos que farão o programa definir o número de grupos ( $K$ ). Este modelo de análise assume que, dentro destes grupos, os locos se encontram em EHW e equilíbrio de ligação (PRITCHARD *et al.*, 2000).

Através deste programa, baseado em frequências genotípicas, os indivíduos podem ser alocados em populações, determinando a ausência ou presença de estrutura populacional, dentre outras utilidades (PRITCHARD *et al.*, 2004; CHEVITARESE, 2009). O valor mais confiável para esta determinação é estimado através de um programa complementar (Structure Harvester), que se baseia em dados de valor de  $\ln PD$  (logaritmo da



probabilidade dos dados) e do desvio padrão encontrados durante a análise estatística (EARL *et al*, 2014).

#### 2.4.3.5 Análise de Agrupamento

A análise de agrupamento é uma maneira de se obter grupos homogêneos, por um esquema que possibilite reunir os indivíduos em um determinado número de grupos, de modo que exista grande homogeneidade dentro de cada grupo e heterogeneidade entre eles (JOHNSON e WICHERN, 1992;).

Dentre as diversas possibilidades, nos métodos hierárquicos de agrupamento, os indivíduos são agrupados por um processo que se repete em vários níveis até que seja estabelecido o dendrograma (diagrama bidimensional em forma de árvore). O dendrograma pode ser considerado a representação simplificada da matriz de dissimilaridade.

Além de várias interpretações paralelas decorrentes desta representação, os níveis de agregação podem ser considerados muito importantes. A posição relativa dos nós costuma ter mais relevância que os valores reais de distâncias entre eles (TENDEIRO, 2005).

As aplicações dos dendrogramas incluem a comparação de populações. Na análise filogenética, pressupõe-se a existência de relação filogenética entre os grupos, sendo que marcadores e análises específicos são requeridos. Na comparação de populações os dendrogramas servem como uma representação gráfica da diversidade dos variantes analisados naquelas populações. Neste cenário, os nós representam similaridade genética porém não significam compartilhamento de ancestralidade, e os clados podem ser formados por diversos motivos, como deriva, migração e outros fatores.

#### 2.4.3.6 Análise de Coordenadas (PCoA) e Componentes Principais (PCA)

As dificuldades encontradas para se interpretar matrizes de distância genética envolvendo muitas populações trouxeram a necessidade da disponibilização de uma forma prática para a visualização dos padrões essenciais das relações genéticas. A análise de coordenadas principais (PCoA)

e a análise de componentes principais (PCA) consistem em ferramentas criadas a partir do mesmo princípio, com o propósito de sintetizar diversas variáveis em menos eixos.

PCoA é um método que permite que se encontre e marque os padrões mais significativos em um conjunto de dados multivariados, como, por exemplo, com múltiplos locos e múltiplas amostras, tanto para similaridades quanto para dissimilaridades. A interligação dos elementos de matriz de distância é identificada baseando-se em suas coordenadas principais (habitualmente três), representadas por combinações de eixos (eixo 1 vs eixo 2, eixo 2 vs eixo 3, eixo 1 vs eixo 3), representados graficamente. Ocasionalmente, é possível plotar as três coordenadas em um gráfico tridimensional. Cada eixo sucessivo explica proporcionalmente menos do total de variação e os primeiros dois ou três eixos revelam tipicamente a maior parte da separação entre grupos. É também possível que a terceira coordenada não seja suficientemente informativa para ser considerada. O espaço dimensional é dividido em quatro quadrantes nos quais as populações, representadas por pontos, que se concentrarem no mesmo quadrante, são consideradas mais semelhantes. Pontos próximos também indicam proximidade genética maior do que pontos distantes.

A análise de componentes principais (PCA) consiste em um método estatístico de análise que realiza uma análise de covariância entre fatores e, desta forma, reduz a dimensionalidade dos dados. A meta é explicar a maior quantidade de variância com o menor número de componentes principais. O objetivo desta transformação é proporcionar ao primeiro componente principal uma variância tão alta quanto possível. Cada componente sucessor será definido para a maior variância possível e não relacionada à dos componentes anteriores (ZARE CHAHOUKI, 2011). A quantidade de variância que os componentes principais explicam é fornecida pela proporção acumulada de cada um. Sugere-se que se retenha os componentes principais que explicam um nível considerado aceitável de variância, o qual dependerá de cada situação. Para fins descritivos, podem ser necessários apenas 80% de variância explicada e, para tanto, habitualmente os primeiros dois a quatro componentes principais costumam ser suficientes para representar a maior parte da variabilidade genética de um conjunto de dados. A representação

ocorre no formato de um diagrama cartesiano no qual os dois primeiros componentes principais seriam as coordenadas e no qual a população é representada por um ponto. Populações geneticamente mais similares tenderão a estar mais próximas nesse espaço bidimensional. Essa representação frequentemente se assemelha ao mapa geográfico, uma vez que populações geograficamente mais próximas tendem a ser geneticamente mais semelhantes (CHEVITARESE, 2009).

As matrizes de distâncias de similaridade geradas para cada uma das análises (PCoA ou PCA) podem ser construídas utilizando métodos diferentes (como Euclidiano e outros), o que pode, entretanto, afetar os resultados de forma significativa. Para conjuntos de dados que incluam valores nulos (como por exemplo falhas de detecção decorrentes de amplificação deficiente para um marcador), como os obtidos para a maioria dos dados de genotipagem usando métodos de *fingerprint*, é aconselhado usar a semelhança de Bray-Curtis e PCoA. A distância de Bray-Curtis não é afetada pelo número de valores nulos entre amostras, como ocorre com a distância euclidiana. O PCA é indicado para conjuntos de dados nos quais não sejam incluídos valores nulos, como é o caso de variáveis ambientais. Isto significa que não é possível usar PCA em caso de dados faltantes.

De maneira geral, PCA limita-se à distância euclidiana, enquanto que PCoA pode utilizar qualquer distância de semelhança desejada. Porém, caso seja possível utilizar a distância euclidiana, ambas as análises (PCoA e PCA) fornecem resultados similares (ZARE CHAHOUKI, 2011). Na presente análise, envolvendo apenas um conjunto de variáveis genéticas, é importante salientar que tanto PCoA quanto PCA não refletem necessariamente ancestralidade comum, apenas um retrato da similaridade das variáveis comparadas.

Os programas estatísticos disponíveis para a execução de tais análises requerem dados muito específicos, tais como frequências genotípicas (para PCoA) ou alélicas para (PCA), de forma que o formato dos dados disponibilizados também é um fator que precisa ser levado em consideração para a escolha da análise a ser conduzida.

## 2.5 PARÂMETROS ESTATÍSTICOS PARA ANÁLISE FORENSE

Como já mencionado, a utilidade e a adequação de marcadores do tipo STR para uso na genética forense requerem que sejam suficientemente discriminativos de forma a oferecerem resultados confiáveis tanto para exames de verossimilhança quanto para exames de paternidade.

Estudos populacionais envolvem, portanto, a determinação de certos parâmetros estatísticos de aplicação forense, tais como heterozigosidade, diversidade gênica, poder de discriminação (*PD*, do inglês, *power of discrimination*), conteúdo de informação polimórfica (*PIC*, do inglês, *polymorphic information content*), probabilidade de coincidência (*PC* ou *MP*, do inglês, *match probability*), poder de exclusão (*PE*) e índice de paternidade (*IP*). Estes parâmetros são determinados a partir de dados estatísticos descritivos, como frequências alélicas, genotípicas, heterozigosidade e homozigosidade.

A homozigosidade (*Ho*) é obtida pelo número de homozigotos observado na amostra, aplicando-se a fórmula a seguir:

$$PC = \sum_{i=1}^n x_i^2$$

(onde *n* = número total de genótipos; *x<sub>i</sub>* = frequência do genótipo *x*)

Heterozigosidade traduz diretamente a proporção de heterozigotos na população em estudo, sendo também referida como heterozigosidade observada. É obtida pela contagem direta de heterozigotos e dividindo-se o valor obtido pelo número total de amostras (indivíduos). Quando presente de forma elevada significa que há maior diversidade alélica, o que reduz a possibilidade de uma coincidência aleatória. Em uma população, a variabilidade genética pode ser quantificada através da heterozigosidade média esperada por loco. Este valor é obtido através da média aritmética dos valores de heterozigosidade de cada um dos vários locos analisados. A somatória dos homozigotos e dos heterozigotos equivale a 100% da amostra (*He* + *Ho* = 1). Logo, o valor de heterozigosidade observada é *He* = 1 – *Ho*.

O termo diversidade gênica é frequentemente utilizado como heterozigosidade esperada (*H*) e reflete a probabilidade de dois alelos escolhidos ao acaso na população serem diferentes. Sua estimativa é realizada através da fórmula a seguir:

$$H \frac{n}{(n-1)} \left[ 1 - \sum_{j=1}^k \left( \frac{n_j}{n} \right)^2 \right] = \frac{n}{(n-1)} \left[ 1 - \sum_{j=1}^k (p_j)^2 \right]$$

(onde  $n$  = número total de genótipos;  $n_j$  = contagem de  $J$  alelos em um loco; e  $p_j$  = frequência alélica)

*Heterozigosidade esperada* difere da *heterozigosidade observada* por se tratar de uma previsão, baseada na frequência alélica conhecida em uma população.

O poder de discriminação (PD) de um loco reflete a probabilidade de dois indivíduos não aparentados e aleatoriamente escolhidos em uma população apresentarem genótipos diferentes entre si, para um marcador específico. Como se trata de uma probabilidade, seu valor máximo é 1. É obtido pela fórmula:

$$PD = 1 - 2(\sum_{i=1}^n p_i^2)^2 + \sum_{i=1}^n p_i^4$$

(onde  $n$  = número total de genótipos;  $p_i$  = frequência alélica de  $i$  alelos em um loco)

A probabilidade de coincidência (PC), também conhecida como probabilidade de identidade, traduz a probabilidade de dois indivíduos escolhidos ao acaso na população possuírem genótipos coincidentes, isto é, idênticos, em um determinado loco. Este termo é o oposto à probabilidade de discriminação descrito no parágrafo anterior. A soma destes valores (PD e PC) é igual a 1. Logo,  $PC = 1 - PD$ .

Para se avaliar a eficiência do sistema completo de marcadores de uso forense, são considerados os valores de discriminação combinados de todos os marcadores usados. Portanto, quanto maior o número de marcadores empregados, maior será a capacidade discriminatória dos sistemas utilizados. Por este motivo, sugere-se o uso mínimo dos 20 marcadores CODIS, numerosos, altamente polimórficos e em equilíbrio de ligação (RODOVALHO *et al.*, 2017) cujos valores podem ser agrupados gerando o poder de discriminação combinado. O valor combinado do poder de discriminação é calculado através da seguinte fórmula:

$$PD_{Comb} = 1 - \prod_{l=1}^L (1 - PD_l)$$

(onde L = número de locos)

O parâmetro denominado conteúdo de informação polimórfica (PIC) representa a informatividade ou capacidade informativa que um determinado polimorfismo confere ao marcador para que seja possível a dedução do parentesco genético. Isto é, indica o nível de polimorfismo de um loco. É obtida através da fórmula a seguir:

$$PIC = 1 - \sum_{i=1}^n p_i^2 - (\sum_{i=1}^n p_i^2)^2 + \sum_{i=1}^n p_i^4$$

(onde n = número total de genótipos;  $p_i$  = frequência alélica de  $i$  alelos em um loco)

O poder de exclusão (PE) *a priori* configura a probabilidade que um marcador apresenta para excluir um indivíduo, sendo reflexo direto do seu grau de polimorfismo. Isto é, quanto mais polimórfico e quanto mais equilibrada se encontrar a distribuição de suas frequências alélicas, tanto maior será a probabilidade de exclusão *a priori*. Estima-se o PE através da seguinte fórmula:

$$PE = 1 - 2\sum_{i=1}^n p_i^2 - 2[\sum_{i=1}^n p_i^2]^2 + 2\sum_{i=1}^n p_i^4 - \sum_{i=1}^n p_i^3 + 3\sum_{i=1}^n p_i^3 + 3\sum_{i=1}^n p_i^2 \sum_{i=1}^n p_i^3$$

(onde n = número total de genótipos;  $p_i$  = frequência alélica de  $i$  alelos em um loco)

O poder de exclusão do sistema completo de marcadores forenses se dá pela combinação do poder de exclusão para cada loco individual, calculado através da heterozigosidade de cada loco (TRACEY, 2001), gerando o poder de exclusão combinado:

$$PE_{Comb} = 1 - \prod_{l=1}^L (1 - PE_l)$$

(onde L = número de locos)

O índice de paternidade (IP) indica quantas vezes é mais provável que um indivíduo, sob investigação, seja realmente o pai biológico por comparação do que um indivíduo aleatoriamente escolhido da população, sendo obtido pela fórmula abaixo:

$$PI = \frac{H+h}{2h} = \frac{(1-h)+h}{2h} = \frac{1}{2h} = \frac{1}{2\sum_{i=1}^n p_i^2}$$

(onde H = heterozigosidade; h = homozigosidade; n = número total de genótipos;  $p_i$  = frequência alélica de  $i$  alelos em um loco)

Em suma, parâmetros de PC, PD e PIC são empregados com objetivo de avaliar a conformidade do marcador para a identificação forense de indivíduos. Os parâmetros de PE e IP são empregados visando a adequação do marcador com relação à investigação de vínculo genético de paternidade.

Idealmente, para fins forenses, para que os marcadores sejam considerados suficientemente informativos, devem apresentar heterozigosidade observada próxima da heterozigosidade esperada e apresentar valores baixos para PC (MP) e elevados para PD, PIC, PE e IP (JOBBLING, 2004; GROVER *et al.*, 2014).



### 3 OBJETIVOS

A caracterização da diversidade populacional baseada em STRs autossômicos de populações nativas sul-americanas, para testes forenses, tem sido limitada a algumas poucas populações e, destas, são poucos os marcadores STR incluídos nos bancos de dados, como o CODIS. Como as frequências genotípicas e alélicas de marcadores em populações isoladas, como as ameríndias, em geral diferem das estimadas para populações urbanas e representadas no banco de dados nacional, a estimativa de dados populacionais para ameríndios é essencial para os cálculos de probabilidades de correspondência aleatória, a qual é ajustada de acordo com a estrutura da população. Sendo assim, os objetivos deste trabalho são:

#### 3.1 OBJETIVO GERAL

Caracterizar a diversidade genética de STRs autossômicos de uso forense em populações ameríndias Guarani e Kaingang do Paraná e do Mato Grosso do Sul, especificamente aqueles que compõem a maior parte do banco de dados CODIS, que permitam a estimativa de um fator de correção  $\theta$  preciso para estas populações.

#### 3.2 OBJETIVOS ESPECÍFICOS

- i) Estimar as frequências alélicas de 21 marcadores STR/CODIS para as populações indígenas Guarani (Guarani M'byá, Guarani Ñandeva e Guarani Kaiowá) e Kaingang (Rio das Cobras e do Rio Ivaí) dos estados do Paraná e Mato Grosso do Sul;
- ii) Investigar a existência de desequilíbrio de ligação entre marcadores do painel de 21 STR/CODIS para cada uma das populações ameríndias aqui analisadas;
- iii) Quantificar a diferenciação genética ( $R_{ST}$  e ou  $F_{ST}$ ) das populações de estudo entre si, destas com outras populações indígenas brasileiras e

ameríndias já estudadas e também com relação a populações urbanas brasileiras e populações de origem predominante europeia, asiática e africana.

- iv) Verificar, através de medidas de diferenciação genética se as três populações Guarani e as duas populações Kaingang apresentam similaridade genética que justifique que seus dados sejam unificados nos estudos forenses;
- v) Estimar parâmetros estatísticos de interesse forense, tais como números de alelos, heterozigosidades observada e esperada, poder de discriminação (PD), poder de exclusão (PE), probabilidade de coincidência aleatória (PM), conteúdo de informação polimórfica (PIC), índice de paternidade (IP), poder de discriminação combinado (PDC) e poder de exclusão combinado (PEC);
- vi) Adicionar os dados aos bancos de frequências alélicas de STRs forenses, os quais carecem de dados de populações ameríndias.

## 4 MÉTODOS

### 4.1 AMOSTRAS POPULACIONAIS

Nesse trabalho, foram estudadas populações de duas reservas indígenas do Paraná: (1) reserva de Rio das Cobras (localizada no município de Nova Laranjeiras, contendo cerca de 1600 indivíduos, dos quais um terço é Kaingang e os demais Guarani ou Xetá); (2) reserva do Ivaí (localizada no município de Manoel Ribas, contendo cerca de 700 Kaingangs). Estas populações vivem em vilas diferentes e passam doravante a ser designadas como KIV (Kaingang do Ivaí), KRC (Kaingang do Rio das Cobras) e GRC (Guarani do Rio das Cobras). A maioria dos indivíduos que compõem estas populações são parentes, de forma que não foi possível selecionar subgrupos de indivíduos completamente não aparentados.

No Mato Grosso do Sul foram selecionadas as comunidades Guarani Kaiowá e Guarani Nandeva, doravante denominadas, respectivamente, GKW e GND. Foram coletadas amostras de Guarani Kaiowá (das aldeias Amambai e Limão Verde, localizadas no município de Amambai, às quais pertencem cerca de 4.500 e 460 indivíduos, respectivamente) e Guarani Nandeva (dos municípios de Amambai e Porto Lindo, na qual residem aproximadamente 1.600 pessoas) (BELTRAME, 2008).

Foram analisados 319 indivíduos pertencentes, portanto, às populações indígenas residentes em reservas indígenas localizadas nos municípios de Nova Laranjeiras (PR), Manoel Ribas (PR), Amambaí (MS), Limão Verde (MS) e Porto Lindo (MS). Destes, 121 são Kaingangs (61 de Ivaí e 60 do Rio das Cobras, no PR) e 198 são Guaranis (51 Guarani M'byá de Rio das Cobras, PR, 71 Guarani Nandeva e 76 Guarani Kaiowá, do MS). Estas populações passam doravante a ser designadas como KIV (Kaingang do Ivaí), KRC (Kaingang do Rio das Cobras), GRC (Guarani do Rio das Cobras), GKW (Guarani Kaiowá) e GND (Guarani Nandeva).

A localização geográfica das populações se encontra representada na Figura 2.



**FIGURA 2** - LOCALIZAÇÃO GEOGRÁFICA DAS ÁREAS DE RESERVA ONDE VIVEM AS POPULAÇÕES INDÍGENAS ANALISADAS NESTE ESTUDO.

(1) Ivaí (KIV: Kaingang do Ivaí). (2) Rio das Cobras (KRC e GRC – Kaingang e Guarani do Rio das Cobras). (3) Amambaí e Limão Verde e (4) Porto Lindo (GND e GKW).

Fonte: Adaptado de Oliveira *et al*, 2008.

As coletas de material biológico e extrações de DNA foram realizadas, para projetos anteriores, pela equipe do Laboratório de Genética Molecular Humana da Universidade Federal do Paraná, conforme descrito anteriormente (PETZL-ERLER, 1993). O DNA das amostras foi extraído a partir do sangue total. Foi utilizado o método de fenol-clorofórmio-álcool isoamílico com modificações (MANIATIS *et al*, 1982) para Guarani M'byá e Kaingang e o método de *salting-out* com modificações (LAHIRI *et al*, 1991) para Guarani Kaiowá e Nandeva. A concentração de suas soluções foi ajustada a 20 ng/μL.

Este trabalho foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos do Setor de Ciências Biológicas. Para a utilização do material biológico de indivíduos indígenas, o projeto também foi aprovado pela Fundação Nacional do Índio (FUNAI) e pelos líderes das populações indígenas estudadas. As amostras das populações indígenas foram colhidas anteriormente à resolução CNS 196/96 do Ministério da Saúde, quando ainda não havia obrigatoriedade do consentimento livre e esclarecido por escrito, e seguiram o código de conduta ética e a legislação brasileira no momento da coleta. A pesquisa contou com a aprovação dos líderes comunitários e todos os indivíduos participaram de forma voluntária. Os pesquisadores ainda garantiram o anonimato dos participantes, que foram esclarecidos em relação aos riscos e benefícios da pesquisa antes da coleta de material biológico.

#### 4.2 AMPLIFICAÇÃO E GENOTIPAGEM

Os procedimentos de quantificação, amplificação e genotipagem foram conduzidos nas dependências do Laboratório de Genética Molecular Forense da Gerência de Laboratórios Forenses da Polícia Científica do Estado do Paraná.

As amostras foram quantificadas através do *kit Quantifiler Trio DNA Quantification* em 7500 Fast Real-Time PCR System (*Thermo Fisher Scientific*) a fim de se mensurar a concentração de DNA, bem como para detectar possível presença de degradação e ou de inibidores de amplificação.

Na sequência, as amostras foram ajustadas à concentração de 0,25ng/μL de DNA e os marcadores de interesse foram amplificados através do *kit* comercial de amplificação *GlobalFiler®* (*Thermo Fisher Scientific*), em reação *multiplex* abrangendo 21 locos autossômicos, em termociclador GeneAmp® PCR System 9700 programado de acordo com protocolo fornecido pelos fabricantes do sistema utilizado, apenas ajustando-se o volume final de reação para 6,25μL. Dentre os marcadores contemplados pelo *kit* encontram-se 20 locos STR que compõem o painel de STR do CODIS (*CSF1PO*, *FGA*, *TH01*, *TPOX*, *VWA*, *D3S1358*, *D5S818*, *D7S820*, *D8S1179*, *D13S317*, *D16S539*, *D18S51*, *D21S11*, *D1S1656*, *D2S441*, *D2S1338*, *D10S1248*, *D12S391*, *D19S433* e *D22S1045*), o marcador *SE33* e os marcadores

*Amelogenina*, *Yindel* e *Y-STR (DYS391)*. Todas as reações de amplificação foram devidamente acompanhadas de um controle positivo (DNA 007, fornecido pelo fabricante do *kit*) e de um controle negativo (água ultrapura que acompanha o *kit* de amplificação).

Alíquotas de 1µL das amostras amplificadas foram na sequência diluídas em 9,6µL de formamida *Hi-Di™* e 0,4µL do marcador interno *GeneScan™ 600 Liz™ Size Standard v2.0*, desnaturadas à 95°C em termociclador *GeneAmp PCR System 9700 (Applied Biosystems)*, e, então, submetidas à corrida eletroforética capilar com polímero POP-4, seguida de detecção da fluorescência laser-induzida no analisador automático de DNA ABI PRISM 3500 (*Applied Biosystems*).

A análise individual e comparativa dos eletroferogramas resultantes foi realizada pelo *Genemapper ID-X Software v. 1.4 (Applied Biosystems)* com o uso do *GlobalFiler® Allelic Ladder (Applied Biosystems)* correspondente.

#### 4.4 COMPARAÇÕES COM OUTRAS POPULAÇÕES

Comparações entre as populações indígenas brasileiras com populações miscigenadas de indivíduos oriundos das cinco macrorregiões geográficas brasileiras foram realizadas. Estas amostras foram constituídas de 265 indivíduos da região sul, 208 da região sudeste, 228 da região centro-oeste, 235 da região nordeste e 171 da região norte, totalizando 1107 indivíduos cujas frequências genotípicas estão disponíveis no trabalho de GUSMÃO *et al*, (2016).

Análises populacionais com outras populações ameríndias, europeias e norte-americanas também foram realizadas e limitadas ao tipo de dado disponível – frequências alélicas ou genotípicas.

Foram obtidas e incluídas em análises comparativas as frequências alélicas para as seguintes 38 populações ameríndias: Terena, Apalai, Kairitiana e Surui do Brasil, Guaraní da Argentina, nativos do Gran Chaco da Argentina, Maya da Guatemala, nativos da Guiana Francesa, Ameríndios da Bolívia, Peru, El Salvador e populações ameríndias da América do Norte, incluindo México e Alaska (AMIGO *et al*, 2009; BUDOWLE *et al*, 2002; ROCABADO *et al*, 2009; LOVO-GÓMEZ *et al*, 2007; SALA *et al*, 2010; CARDOSO *et al*, 2016;

KANTHASWAMY *et al*, 2016; BUDOWLE *et al*, 2001; RANGEL-VILLALOBOS *et al*, 2013; SILVA *et al*, 2012; MANTA *et al*, 2012; CALLEGARI-JACQUES *et al*, 2011; MAZIERES *et al*, 2011; CROSSETTI *et al*, 2008).

Foram acrescentadas às análises populações doravante denominadas “parentais”, provenientes de continentes diferentes. Aqui foram selecionadas seis populações, sendo quatro europeias (Espanha, Portugal, Itália e Alemanha), uma africana e uma japonesa (TIE *et al*, 2006; CABEZAS *et al*, 2015, STRIDER, 2019). Foi utilizado também um grupo de populações norte americanas de origem predominante: europeia (EUR), hispânica (HISP), asiática (ASIAN) e afroamericana (AA) (HILL *et al*, 2013). Salienta-se que “hispânicos”, assim como “latinos”, são os termos usados nos Estados Unidos da América para classificar os americanos com origens nos países hispânicos da América Latina ou na Espanha e, em geral, incluem todas as pessoas que se autoclassificam como hispânico ou latino.

As publicações sobre marcadores STR em populações nem sempre disponibilizam dados de frequências genotípicas, apenas as frequências alélicas. Uma vez que parte das análises pretendidas (Estatística R e PCoA) requer o uso de frequências genotípicas, tais análises foram, quando necessário, substituídas pela estatística F e PCA. Para a realização das análises no programa Structure, foram utilizadas somente as populações miscigenadas brasileiras e norte-americanas.

O painel de marcadores STR utilizado difere, em maior ou menor proporção, dentre as diversas publicações populacionais disponíveis. Foram selecionados apenas 6 marcadores (D3S1358, D8S1179, D21S11, D18S51, D19S539 e TH01), coincidentes em todas estas populações, para a realização das análises que envolveram os nativos das Américas e as populações parentais.

#### 4.3 ANÁLISE DOS DADOS

As frequências alélicas e genotípicas foram obtidas por contagem direta. A adequação das frequências genotípicas ao equilíbrio de Hardy-Weinberg foi verificada através do algoritmo *random walk*, que utiliza uma versão modificada da cadeia de Markov (GUO e THOMPSON, 1992) implementado no



pacote de programas ARLEQUIN v. 3.5 (EXCOFFIER *et al.*, 2005). O mesmo pacote de programas foi utilizado para se estimar as frequências esperadas para a associação entre alelos de diferentes locos através do algoritmo ELB. Para os locos em DL ( $p < 0,05$ ), verificou-se a magnitude do desequilíbrio através do cálculo do valor de  $D'$ , através dos pacotes de programas GENEPOP v. 4.2 e PyPopWin32 – 0.7.0, adotando como limite de significância  $p < 0,05$  (RAYMOND *et al.*, 2008; LANCASTER *et al.*, 2003).

A frequência alélica mínima (FAM) foi calculada pela fórmula  $5/2N$  (NRC II, 1996).

Os parâmetros de diferenciação populacional foram gerados a partir das frequências genotípicas (para estatística R), através do programa GenAlex (PEAKALL *et al.*, 2006), e das frequências alélicas (para estatística F), pelo programa POPTREEW (TAKEZAKI *et al.*, 2014).

Os parâmetros forenses de heterozigosidade, poder de discriminação (PD), poder de coincidência (PC ou MP), conteúdo de informação polimórfica (PIC), índice de paternidade (IP) e poder de exclusão (PE) foram estimados com o programa PowerStats v12 (Tereba, 1999). A determinação dos valores de Poder de Discriminação Combinado ( $PD_{COMB}$ ) e Poder de Exclusão Combinado ( $PE_{COMB}$ ) para todo conjunto de 21 marcadores foi executada pelo Microsoft Excel através das fórmulas referidas a seguir, baseadas nos índices individuais para cada parâmetro e para cada loco (SWGDAM, 2010):

$$PD_{COMB} = 1 - [(1-PD_1) \times (1-PD_2) \times \dots \times (1-PD_{21})]$$

$$PE_{COMB} = 1 - [(1-PE_1) \times (1-PE_2) \times \dots \times (1-PE_{21})]$$

A comparação entre as populações foi realizada através de teste exato de diferenciação populacional através da estatística R ( $R_{ST}$  e  $R_{IS}$ ) e da estatística F ( $F_{ST}$ ) par a par. A Análise de Variância Molecular (AMOVA) loco a loco foi utilizada para detectar a distribuição da diversidade genética dentro e entre as populações. Todas estas análises foram realizadas no pacote de programas ARLEQUIN v. 3.5 (EXCOFFIER *et al.*, 2005) e GenAlex (PEAKALL *et al.*, 2006).

Foi realizada a Análise de Coordenadas Principais (PCoA) para as populações para as quais as frequências genotípicas se encontravam disponíveis (populações ameríndias, brasileira miscigenada e norte americana com ancestralidades predominantemente europeia, africana, hispânica e asiática), utilizando o programa GenAlex (PEAKALL *et al*, 2006).

A Análise de Componente Principal (PCA) foi realizada para as populações ameríndias objeto do presente trabalho e as demais populações indígenas das Américas, bem como a populações europeias, japonesas e africanas. Para tal análise, foram usadas as frequências alélicas disponíveis, utilizando o programa Minitab 17 Statistical Software (2010), sendo considerados os componentes com autovalores superiores a 1, segundo o critério de Kaiser-Meyer-Olkin (KMO) de teste de adequação da amostra, de forma a obter no mínimo 80% da variância explicada. Foram utilizados os seis marcadores (D3S1358, D8S1179, D21S11, D18S51, D16S539 e TH01) coincidentes entre todas as populações.

Para inferir o número de grupos (K) nos quais os genótipos se apresentam estruturados, foi utilizado o programa Structure (PRITCHARD *et al*, 2000). A inferência de estrutura populacional foi realizada pela estimativa de proporções individuais de ancestralidade para os indígenas Guarani e Kaingang entre si (k variando de 1 a 5), para os indígenas com relação à população miscigenada brasileira (k variando de 1 a 10) e para os indígenas com relação a populações norte americanas citadas acima (k variando de 1 a 9). Foram utilizadas 3 repetições para cada valor de K, com 100.000 interações iniciais (burn-ins) e 100.000 simulações de Monte Carlo, via Cadeias de Markov (MCMC). Foi selecionado o modelo que assume que há miscigenação entre as populações (*admixture model*) e o modelo que assume frequências alélicas independentes entre as diferentes populações. Para a definição do valor de K mais provável em relação aos propostos, foi utilizado o valor de Delta K fornecido através do programa Structure Harvester (EVANNO *et al*, 2005; EARL *et al*, 2014).

As distâncias genéticas foram utilizadas para a construção de dendrogramas utilizando o método *Neighbor-Joining* (SAITOU *et al*, 1987) no programa POPTREEW (TAKEZAKI *et al*, 2014).

Para a interpretação dos dados estatísticos nas análises que se baseiam no valor de significância ( $p$ ) foi, eventualmente, aplicada a correção de Bonferroni ( $\alpha/N$ ) para múltiplos testes. Cálculos estatísticos que envolvam múltiplas comparações simultâneas podem gerar falsos positivos. Nestes cenários é empregada a correção, que considera que um valor específico de alfa pode ser apropriado para comparações individuais, porém não o seria para o conjunto de todas as comparações. Em tais situações, o valor de alfa deveria ser ajustado (reduzido) de forma a contabilizar o número de comparações que forem executadas, diminuindo a probabilidade do erro tipo I (em que se rejeita a hipótese nula quando ela é, de fato, verdadeira). Na correção de Bonferroni, o valor de alfa *corrigido* é definido para todo o conjunto de comparações, sendo o valor de alfa, originalmente estipulado, dividido pelo número de comparações realizadas ( $\alpha/N$ ) (WEISSTEIN, 2011).

## 5 RESULTADOS E DISCUSSÃO

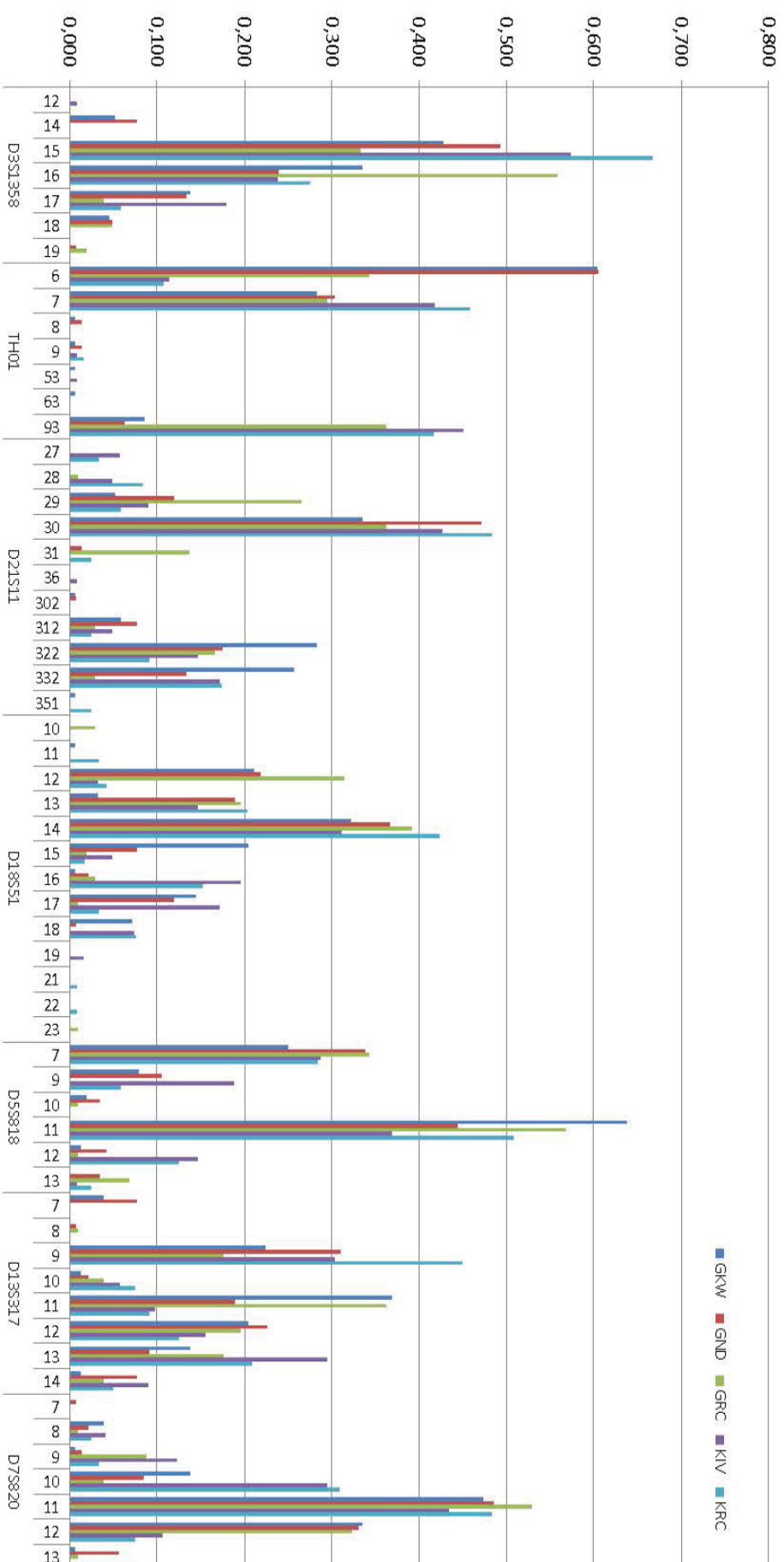
### 5.1 CARACTERIZAÇÃO DA DIVERSIDADE GENÉTICA EM GUARANIS E KAINGANGS: FREQUÊNCIAS ALÉLICAS E EHW

Foram obtidas as frequências genotípicas de vinte e um marcadores STR para as cinco populações ameríndias e, a partir destas, foram estimadas as frequências alélicas, a partir dos genótipos, para cada marcador e para cada população individualmente (Figura 3).

Considerando as cinco populações indígenas como uma população, os vinte e um STRs genotipados amplificaram um total de 186 alelos. Destes, 146 foram genotipados para indivíduos oriundos de GKW, 148 para GND, 118 para GRC, 126 para KIV e 133 para KRC. Em média, foram encontrados 8,8 alelos por marcador estudado. A média total de alelos detectados para cada população Guarani (GKW, GND e GRC) foi de 6,95, 7,05 e 5,62, respectivamente. Entre os Kaingangs, a média de alelos detectados foi de 6 e 6,33, para KIV e KRC, respectivamente.

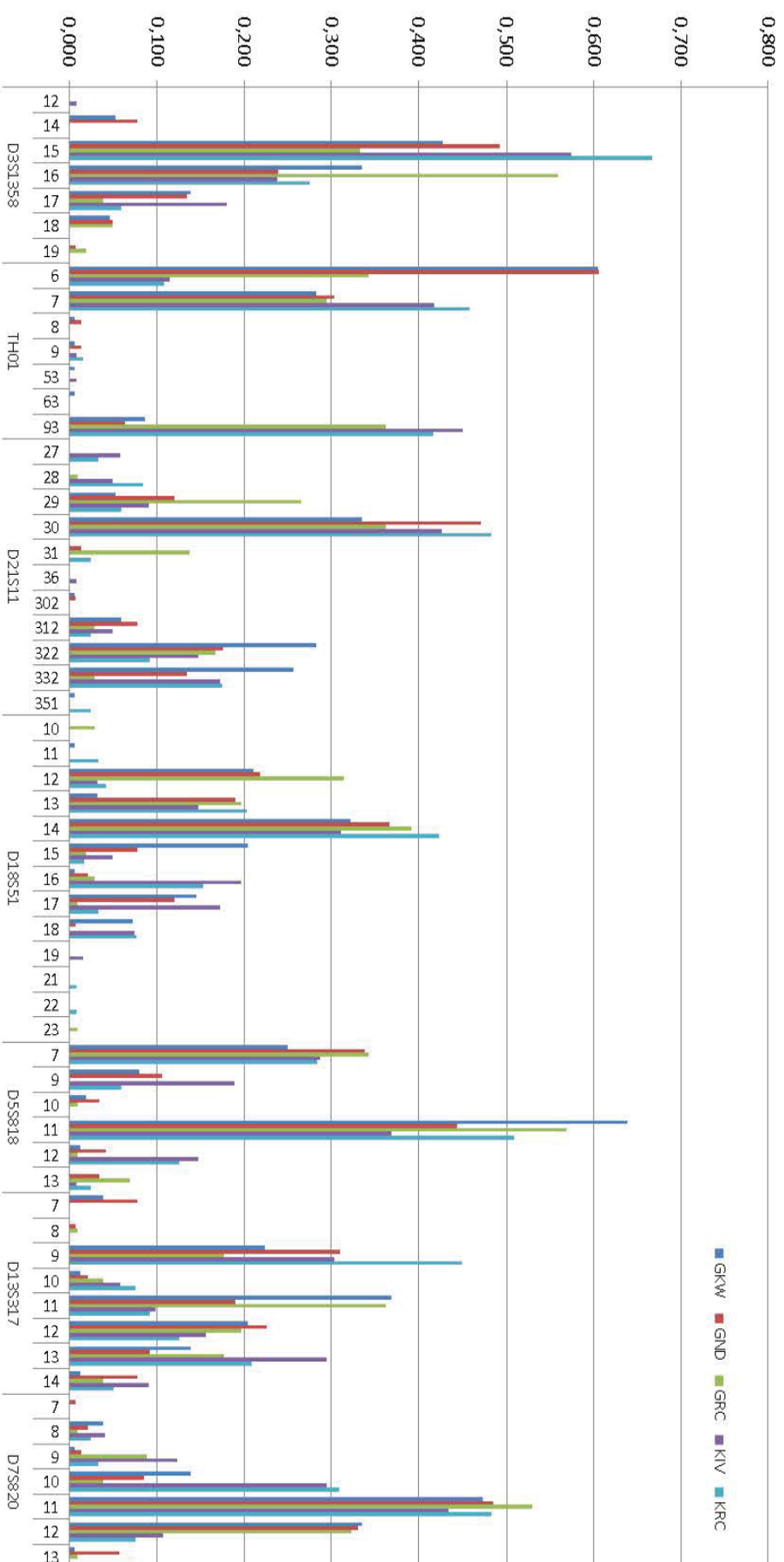
A adequação ao equilíbrio de HW foi verificada para todos os locos após a correção de Bonferroni ( $p_c=0,0004$ ) (Tabela 1). A maior parte das publicações sobre frequências de STRs forenses em populações ameríndias também refere adequação ao equilíbrio de HW, quando não para todos, para a maioria dos marcadores (KANTHASWAMY *et al*, 2016; RANGEL-VILLALOBOS *et al*, 2013; BUDOWLE *et al*, 2002; LOVO-GÓMEZ *et al*, 2006; SALA *et al*, 2010; CARDOSO *et al*, 2016; MANTA *et al*, 2012; MAZIERES *et al*, 2011; CROSSETI *et al*, 2008).

O marcador com menor número de alelos nas populações ameríndias foi CSF1PO, com média de 4,2 alelos. O marcador SE33 apresentou o maior número de alelos, com média de 13 alelos presentes entre os indígenas.



**FIGURA 3 - FREQUÊNCIAS ALÉLICAS DE 21 MARCADORES STR NAS 5 POPULAÇÕES AMERÍNDIAS**

(continua)

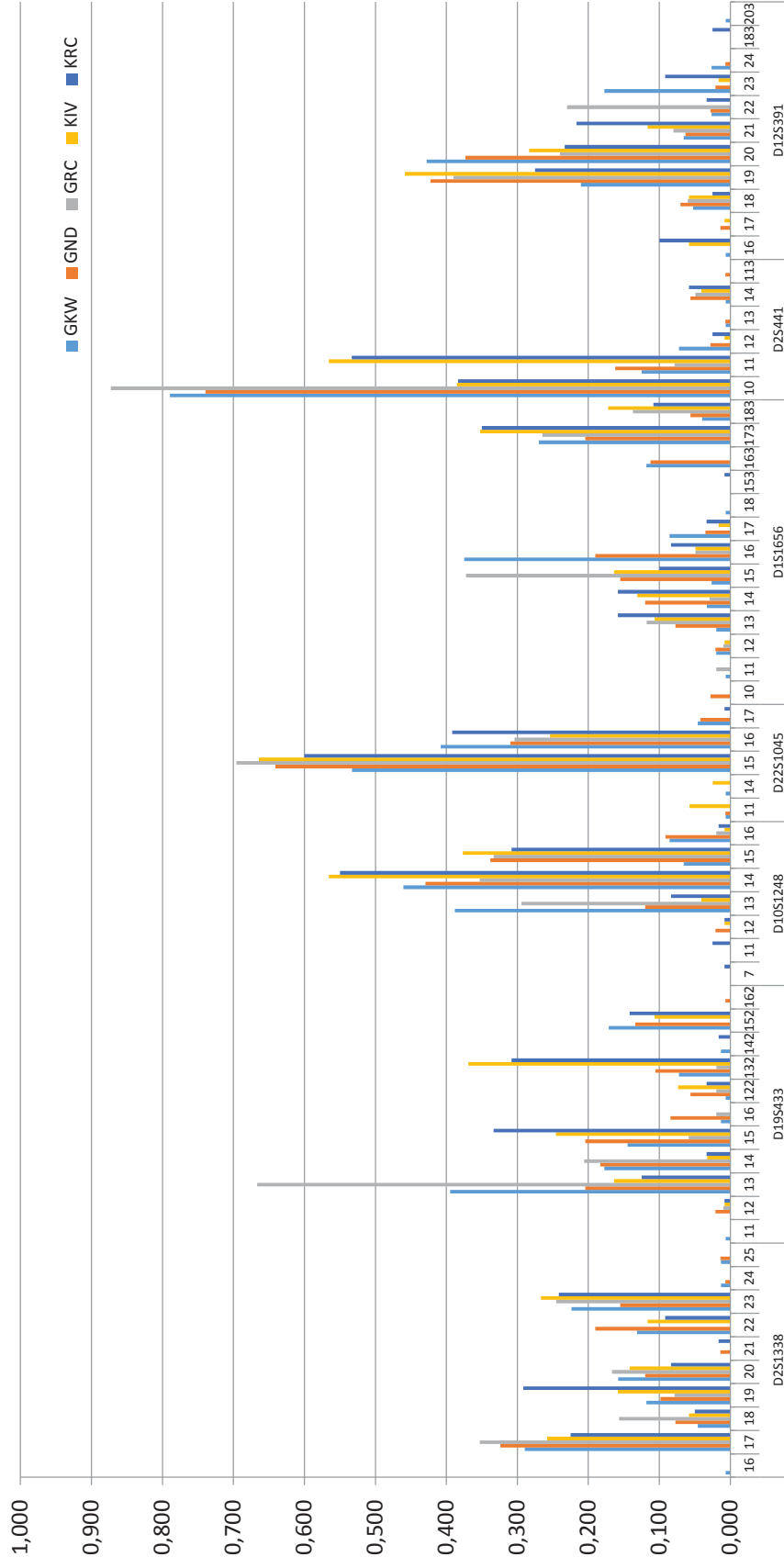


**FIGURA 3 - FREQUÊNCIAS ALÉLICAS DE 21 MARCADORES STR NAS 5 POPULAÇÕES AMERÍNDIAS**

(continuação)







**FIGURA 3 - FREQUÊNCIAS ALÉLICAS DE 21 MARCADORES STR NAS 5 POPULAÇÕES AMERÍNDIAS**

FONTE: O Autor (2019).

(conclusão)

**TABELA 1** – VALOR DE *P* REFERENTE AO EQUILÍBRIO DE HARDY-WEINBERG PARA CADA MARCADOR POR POPULAÇÃO INDÍGENA.

Loco	GKW	GND	GRC	KIV	KRC
D3S1358	0,2346	0,2632	0,1145	0,2228	0,9304
TH01	0,0011	0,6106	0,6040	0,0849	0,1170
D21S11	0,5595	0,9667	0,8223	0,4353	0,5483
D18S51	0,5893	0,4215	0,0222	0,0331	0,6971
D5S818	0,2231	0,2304	0,0911	0,0017	0,2033
D13S317	0,5522	0,9480	0,2730	0,4327	0,2104
D7S820	0,0529	0,0137	0,5932	0,2359	0,0037
D16S539	0,2643	0,0092	0,4549	0,1815	0,6425
CSF1PO	0,2853	0,1504	0,3175	0,5153	0,4906
SE33	0,5262	0,1049	0,0436	0,1232	0,1409
vWA	0,7770	0,0087	0,8627	0,2753	0,4817
D8S1179	0,6471	0,1291	0,1038	0,0607	0,7813
TPOX	0,6431	0,7901	0,4457	0,4882	0,1901
FGA	0,4890	0,3855	0,6006	0,3587	0,1571
D2S1338	0,0821	0,3605	0,9962	0,3535	0,1294
D19S433	0,9955	0,1360	0,0715	0,6221	0,2348
D10S1248	0,2633	0,6820	0,7632	1	0,9358
D22S1045	0,1589	0,4215	0,1049	0,7826	0,4786
D1S1656	0,3293	0,4882	0,1765	0,0229	0,2344
D2S441	1	0,9502	0,0363	0,6128	0,9296
D12S391	0,0633	0,2073	0,2464	0,7461	0,0381

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. Nível de significância  $p < 0,0004$  (após correção de Bonferroni).

FONTE: O autor (2019).

As frequências alélicas mais altas foram observadas, em sua maioria, para alelos de sequência repetitiva completa, isto é, aqueles cujo motivo repetitivo apresenta exatamente os mesmos nucleotídeos em todas as repetições. Os microvariantes (que apresentam um motivo incompleto) mais frequentes foram TH01\*9.3 (frequências variando de 0,0638 a 0,4508), D1S1656\*17.3 (0,2042 a 0,3524), D19S433\*13.2 (0,0196 a 0,3688) e SE33\*27.2 (0,0294 a 0,225).

Diversos dos alelos mais frequentes, para um determinado marcador, apresentam, entre si, a diferença de uma unidade de repetição, como é o caso

dos alelos D3S1358\*15 e D3S1358\*16. Repetidamente os alelos mais frequentes em uma população também o são em outras.

Para cada marcador, identificou-se aqueles alelos mais frequentes. Em alguns marcadores, de todos os alelos detectados, dois ou três alelos coincidem de ser os mais frequentes entre as cinco populações, como, por exemplo, os alelos D5S818\*7 e D5S818\*11.

Foi possível detectar que, para seis marcadores, os alelos mais frequentes coincidem para a população indígena total: D21S11\*30, D18S51\*14, D5S818\*11, D7S820\*11, D10S1248\*14 e D22S1045\*15.

Os alelos vWA\*16, D8S1179\*14, TPOX\*11, D2S1338\*17, D19S433\*13 e D2S441\*10 foram os mais frequentes quando consideradas as três populações Guaranis. Kaingangs apresentaram, como os mais frequentes, os alelos D8S1179\*14, TPOX\*8, D2S441\*11 e vWA\*17.

Foram detectadas diferenças entre frequências de alguns alelos compartilhados por todas as populações indígenas, quando estas foram comparadas entre si. Por exemplo, entre os ameríndios, o alelo D22S391\*22 é 11 vezes mais frequente em GRC do que os outros Guaranis e 7 vezes mais frequente do que nos KRC, não tendo sido detectado nos KIV. O alelo D2S441\*11 é 7 vezes menos frequente nos GRC do que nos Kaingangs e apresenta metade da frequência dos demais Guaranis. Os alelos D10S1248\*13 e D19S433\*13 são mais frequentes entre Guaranis.

As diferenças das frequências de alelos compartilhados pelas populações indígenas entre si poderiam se dever à baixa ocorrência de fluxo gênico entre os ameríndios, o qual já foi descrito como sendo bastante reduzido (1,4% dos Guaranis apresentam similaridade genética com Kaingangs e 0,5% dos Kaingangs apresentam similaridade genética com Guaranis) em razão das tradições mutuamente segregantes, como culturais e, principalmente, linguísticas, presentes nestas comunidades (PETZL-ERLER *et al*, 1993).

As frequências dos vinte e um marcadores STR, como a distribuição das frequências alélicas observadas para as populações indígenas Guarani (GKW, GND e GRC) e Kaingang (KIV e KRC) do Paraná e Mato Grosso do Sul encontram-se incluídas no Anexo I – Tabelas 16 a 20.

Foram encontrados alelos com frequências inferiores à frequência alélica mínima (FAM) calculada a partir da fórmula  $5/2Nm$  em todas as populações ameríndias Guaranis e Kaingangs estudadas no presente trabalho. Estes alelos são considerados raros. Para que análises forenses baseadas em frequências alélicas que incluam alelos raros sejam o mais fidedignas possível, estes devem ser observados em um mínimo de cinco indivíduos diferentes em uma determinada população (NRC II, 1996; HESSAB *et al*, 2018). Ao mesmo tempo, contemporizando a possível ocorrência de viés nos cálculos estatísticos, convencionou-se, que a frequência observada seja substituída pela chamada frequência alélica mínima (FAM), cujo valor pode ser considerado um limiar conservador diretamente influenciado pelo tamanho e pela heterozigosidade amostral (BUDOWLE, *et al*, 1996; HESSAB *et al*, 2018). A FAM deve ser calculada para cada marcador e população individualmente, segundo a fórmula  $5/2N$  (NRC II, 1996). Este critério vem, por conseguinte, sendo adotado em publicações com dados de frequências alélicas com propósitos forenses (CABEZAS SILVA *et al*, 2016). Dessa forma, para a elaboração de banco de frequências alélicas das populações indígenas testadas, recomenda-se que seja feito um ajuste para os alelos raros detectados, substituindo-se a frequência observada, quando inferior a  $5/2N$ , pela FAM. O valor individual da FAM foi calculado para cada marcador e cada população ameríndia do presente trabalho, encontrando-se no Anexo II – Tabela 21.

Os alelos encontrados, de maneira geral, estão presentes, em maior ou menor frequência, na maioria das cinco populações indígenas. Entretanto, na comparação das cinco populações ameríndias entre si, foi constatada a ocorrência de trinta e três alelos exclusivos, isto é, que aparecem somente uma vez, para uma das cinco populações ameríndias. Destes, GKW, GND e KRC apresentam, cada qual, oito; cinco alelos exclusivos foram encontrados em GRC; e quatro foram constatados para KIV (Tabela 2). A amplitude da variação das frequências dos alelos exclusivos detectados nos ameríndios variou de 0,007 a 0,186 (média de 0,021, indicando relativa baixa frequência). O alelo exclusivo com a maior frequência observada foi vWA\*15, detectado apenas dentre GRC. O único alelo que permaneceu exclusivo das populações indígenas foi D16S539\*15, encontrado somente na população GRC.

**TABELA 2 - ALELOS EXCLUSIVOS DETECTADOS NA COMPARAÇÃO DAS FREQUÊNCIAS ALÉLICAS DOS AMERÍNDIOS ENTRE SI.**

População	Marcador	Alelo exclusivo	Frequência
GKW	TH01	6.3	0,007
	CSF1PO	9	0,007
	SE33	31.2	0,046
	vWA	20	0,007
	D2S1338	16	0,007
	D19S433	11	0,007
	D1S1656	18	0,007
	D12S391	20.3	0,007
GND	D7S820	7	0,007
	SE33	13	0,014
	D8S1179	8	0,014
	D8S1179	16	0,014
	FGA	18	0,007
	D19S433	16.2	0,007
	D1S1656	10	0,028
	D2S441	11.3	0,007
GRC	D18S51	10	0,029
	D18S51	23	0,010
	<b>D16S539</b>	<b>15</b>	<b>0,029**</b>
	SE33	28	0,108
	<b>vWA</b>	<b>15</b>	<b>0,186*</b>
KIV	D3S1358	12	0,008
	D21S11	36	0,008
	D18S51	19	0,016
	FGA	43.2	0,016
KRC	D18S51	21	0,008
	D18S51	22	0,008
	SE33	24.2	0,008
	TPOX	5	0,008
	D10S1248	7	0,008
	D10S1248	11	0,025
	D1S1656	15.3	0,008
	D12S391	18.3	0,025

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. \*alelo privado indígena em maior frequência. \*\*Alelo exclusivo remanescente após comparação com população brasileira.

FONTE: O autor (2019).

Segundo Wang *et al* (2007), os povos ameríndios apresentam reduzida quantidade de alelos exclusivos e, se considerarmos apenas os povos indígenas do sudeste americano, esta ocorrência seria ainda menor. Além disso, as frequências dos alelos exclusivos costumam, de maneira geral, ser baixas nestas populações, como observado para os indígenas em estudo.

#### 5.1.2 Desequilíbrio de ligação

Na avaliação dos locos com relação à presença de desequilíbrio de ligação, após a correção de Bonferroni ( $p_c=0,0004$ ), das 210 comparações par a par realizadas, foi detectada associação entre 22 pares de locos nas cinco populações indígenas (3 pares em GND, 9 pares em KRC, 5 pares em KIV e 5 pares em KRC) (Tabela 3).

Foi estimada a magnitude do  $DL$ , através da determinação do parâmetro  $D'$  (que pode variar de 0 a 1), o qual se mostrou baixo para todos os marcadores com significância estatística para o  $DL$ .

Para a população ameríndia do México, foram igualmente relatados locos STR autossômicos com  $DL$ , dos quais dois pares coincidem com os encontrados para os ameríndios Guarani e Kaingang (D3S1358/D2S1338 e D2S1338/FGA) (RANGEL-VILLALOBOS *et al*, 2013). O resultado obtido para as cinco populações indígenas difere do verificado em populações urbanas brasileiras, nas quais não foram verificadas, após a aplicação de correção de Bonferroni, ocorrências significativas de  $DL$  (GUSMÃO *et al*, 2016; LEITE *et al*, 2009).

Desequilíbrio de ligação nesta configuração pode se dever a cenários envolvendo efeitos demográficos, como existência de subestrutura populacional, deriva e fluxo gênico entre populações adjacentes, gerando efeito de Wahlund entre as comunidades indígenas (redução de heterozigotos) ou alta endogamia decorrente do isolamento e do reduzido número populacional (RANGEL-VILLALOBOS *et al*, 2013; LEITE *et al*, 2009). Em se tratando de populações isoladas e de tamanho finito e reduzido, as populações ameríndias Kaingang e Guarani se enquadram nestas possíveis dinâmicas populacionais (RANGEL-VILLALOBOS *et al*, 2013).

Inferências a respeito de provável estruturação populacional e do tamanho efetivo da população, por exercerem significativa interferência no *DL*, podem se fazer necessárias para que o fenômeno de *DL* possa ser melhor compreendido em uma dada população. Quanto menor o tamanho efetivo da população, por exemplo, tanto mais pronunciado será o *DL* gerado (NORDBORG *et al*, 2008). O *DL* constatado no presente trabalho, entre locos localizados em cromossomos distintos, em populações isoladas, de magnitude pequena e finita, pode apresentar correlação com o tamanho efetivo das populações estudadas.

**TABELA 3** - PARES DE MARCADORES EM DESEQUILÍBRIO DE LIGAÇÃO EM CADA POPULAÇÃO AMERÍNDIA.

População	Par de marcadores	<i>P</i>	<i>D'</i>
<b>GND</b>	TH01:D16S539	$p < 10^{-5}$	0,28355
	D5S818:D16S539	$p < 10^{-5}$	0,31400
<b>GRC</b>	SE33:FGA	$p < 10^{-5}$	0,45750
	D18S51:D13S317	$p < 10^{-5}$	0,46527
	D18S51:SE33	$p < 10^{-5}$	0,50919
	D5S818:SE33	$p < 10^{-5}$	0,38153
	D7S820:SE33	$p < 10^{-5}$	0,49677
	SE33:D8S1179	$p < 10^{-5}$	0,39205
	SE33:FGA	$p < 10^{-5}$	0,50854
	SE33:D2S1338	$p < 10^{-5}$	0,49410
	FGA:D2S1338	$p < 10^{-5}$	0,48469
	SE33:D19S433	$p < 10^{-5}$	0,55682
<b>KIV</b>	D18S51:SE33	$p < 10^{-5}$	0,54970
	D18S51:FGA	$p < 10^{-5}$	0,51728
	D3S1358:D2S1338	$p < 10^{-5}$	0,33959
	SE33:D2S1338	$p < 10^{-5}$	0,50075
<b>KRC</b>	FGA:D22S1045	$p < 10^{-5}$	0,50110
	D13S317:D7S820	$p < 10^{-5}$	0,35348
	TPOX:D10S1248	$p < 10^{-5}$	0,41596
	D7S820:D1S1656	$p < 10^{-5}$	0,38975
	VWA:D2S441	$p < 10^{-5}$	0,45040
	D2S441:D12S391	$p < 10^{-5}$	0,39430

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. Valor de *p* significativo  $< 10^{-5}$ .

FONTE: O autor (2019).



Por esta razão, foi estimada a magnitude do  $DL$ , através da determinação do parâmetro  $D'$ , que varia de 0 a 1. O valor máximo obtido foi de 0,55682, para o par SE33:D19S433, sendo inferior para todos os marcadores com valor de  $p$  significativo para  $DL$ . Além disso, a associação constatada na população ameríndia em estudo não se fez presente entre locos sintênicos, limitando-se a pares localizados em cromossomos distintos. Dessa forma, estes marcadores podem ser considerados não ligados e podem ser tratados como locos independentes, considerando as populações indígenas estudadas, para fins forenses (CABEZAS SILVA *et al*, 2016).

## 5.2 DISTRIBUIÇÃO DA DIVERSIDADE GENÉTICA MOSTRA ESTRUTURAÇÃO POPULACIONAL ENTRE GUARANIS, MAS NÃO ENTRE KAINGANGS

As resultados obtidos pela estatística R e pelas análises realizadas no programa Structure, juntamente com os resultados gerados pelo dendrograma e PCoA, convergem para a indicação de que existe estruturação populacional entre Guaranis (GKW, GND e GRC), porém não entre Kaingangs, como mostrado a seguir.

### 5.2.1 AMOVA e Estatística R

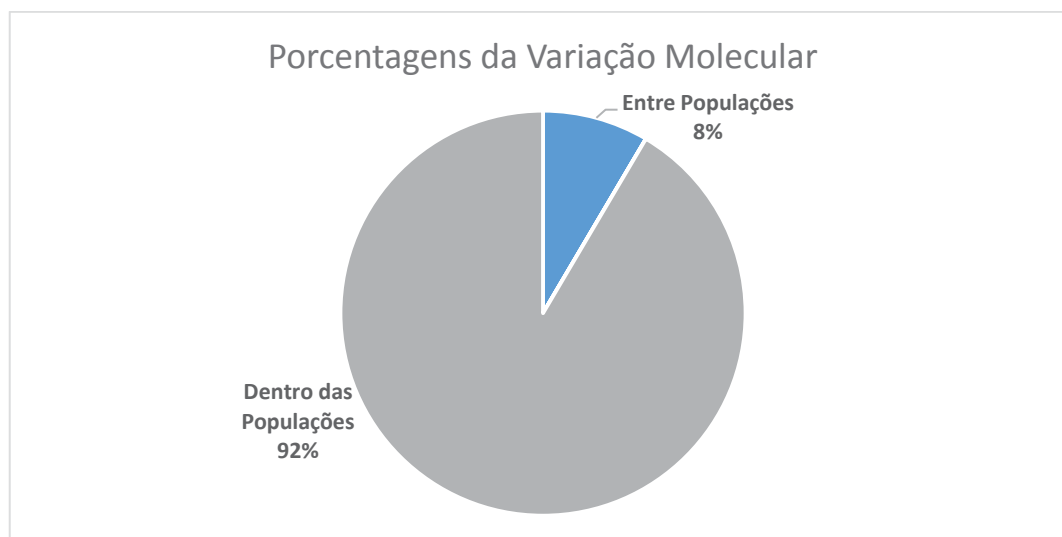
Os valores de  $R_{st}$ ,  $V_a$  (interpopulacional) e  $V_b$  (intrapopulacional) obtidos por AMOVA entre os indígenas, para todos os marcadores testados (Tabela 4) mostram que a variância dentro das populações representou a maior contribuição (81% a 100%) à variância total, isto é, a maior parte da variabilidade genética (92%) está uniformemente distribuída dentro das populações (Figura 4). Após a correção de Bonferroni ( $p < 0,00238$ ), valores de  $R_{st}$  estatisticamente significativos foram encontrados para 10 marcadores: TH01, D21S11, D18S51, D7S820, D16S539, SE33, VWA, D19S433, D10S1248 e D22S1045. A diferenciação genética ( $R_{st}$ ) estimada par a par para estes 11 marcadores entre as cinco populações ameríndias encontra-se no Anexo IV – Tabela 23.

A maior parte da variabilidade genética encontra-se uniformemente distribuída dentro das populações indígenas, como também foi observado em outras publicações sobre STR autossômicos em ameríndios (BURGUETE-ARGUETA *et al*, 2015; GUSMÃO *et al*, 2016; MARTINEZ-GONZALEZ *et al*, 2015).

**TABELA 4** - AMOVA LOCO POR LOCO BASEADO NOS VALORES DE RST CONSIDERANDO AS POPULAÇÕES INDÍGENAS REUNIDAS.

Marcador	Va	% Variação interpopulacional	Vb	% Variação intrapopulacional	Rst	P (para Rst)
<b>D3S1358</b>	0,021	3%	0,762	97%	0,026	0,004
<b>TH01</b>	279,433	19%	1219,918	81%	0,186	0,001*
<b>D21S11</b>	1811,161	8%	19616,511	92%	0,085	0,001*
<b>D18S51</b>	0,425	10%	3,634	90%	0,105	0,001*
<b>D5S818</b>	0,000	0%	3,702	100%	0,000	0,433
<b>D13S317</b>	0,072	2%	2,986	98%	0,024	0,004
<b>D7S820</b>	0,099	10%	0,927	90%	0,097	0,001*
<b>D16S539</b>	0,254	13%	1,630	87%	0,135	0,001*
<b>CSF1PO</b>	0,018	2%	1,041	98%	0,017	0,007
<b>SE33</b>	1708,189	10%	15748,008	90%	0,098	0,001*
<b>VWA</b>	0,113	14%	0,668	86%	0,145	0,001*
<b>D8S1179</b>	0,052	3%	1,824	97%	0,028	0,003
<b>TPOX</b>	0,041	1%	2,696	99%	0,015	0,026
<b>FGA</b>	3,855	1%	524,007	99%	0,007	0,112
<b>D2S1338</b>	0,005	0%	6,939	100%	0,001	0,373
<b>D19S433</b>	551,702	15%	3073,574	85%	0,152	0,001*
<b>D10S1248</b>	0,045	6%	0,756	94%	0,056	0,001*
<b>D22S1045</b>	0,031	5%	0,533	95%	0,055	0,001*
<b>D1S1656</b>	57,413	1%	6217,230	99%	0,009	0,089
<b>D2S441</b>	0,016	0%	17,267	100%	0,001	0,303
<b>D12S391</b>	2,439	1%	179,759	99%	0,013	0,020

Va=variância entre populações; Vb=variâncias individuais dentro da população; \*p<0,0024.



**FIGURA 4** - COMPONENTES DE VARIAÇÃO ENTRE POPULAÇÕES E DENTRO DAS POPULAÇÕES, PARA TODOS OS INDÍGENAS GUARANI (GKW, GND E GRC) E KAINGANG (KIV E KRC).

FONTE: O Autor (2019)

Entre as três populações Guaraní a diferenciação ( $R_{st}$ ) varia de baixa, com relação ao par GKW e GND ( $R_{st}=0,0402$ ); média, com relação ao par GRC e GND ( $R_{st}=0,0554$ ); a alta, com relação ao par GKW e GRC ( $R_{st}=0,1557$ ) (Tabela 5).

As populações Kaingang apresentaram diferenciação muito baixa ( $R_{st}=0,0017$ ) entre si (Tabela 5).

KIV apresentou diferenciação média com relação às três populações Guaraní, com maior intensidade ( $R_{st}=0,1114$ ) em relação a GKW. Os valores encontrados para KIV com relação à GND ( $R_{st}=0,056$ ) e à GRC ( $R_{st}=0,059$ ) assemelham-se aos obtidos entre KRC e GRC ( $R_{st}=0,062$ ). Também observou-se diferenciação média entre KRC e GND ( $R_{st}=0,091$ ). Para o par KRC e GKW foi encontrado o maior valor de  $R_{st}$  entre todas as comparações realizadas entre o pool de indígenas ( $R_{st}=0,1604$ ), indicando diferenciação alta.

A diferenciação par a par entre as populações demonstrou que Guaraní se comportam como três populações diferentes, das quais GKW é a mais diferenciada.

**TABELA 5** - VALORES DE DIFERENCIAÇÃO ( $R_{st}$ ) PARA AS 5 POPULAÇÕES INDÍGENAS

	<i>GKW</i>	<i>GND</i>	<i>GRC</i>	<i>KIV</i>	<i>KRC</i>
<i>GKW</i>	-	+	+	+	+
<i>GND</i>	0,0402	-	+	+	+
<i>GRC</i>	0,1557	0,0554	-	+	+
<i>KIV</i>	0,1114	0,0590	0,0599	-	-
<i>KRC</i>	0,1604	0,0912	0,0617	0,0017	-

GKW: Guaraní Kaiowá; GND: Guaraní Nandeva; GRC: Guaraní M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras.  $R_{st}<0,05$  – baixa diferenciação genética.  $0,05<R_{st}<0,15$  – média diferenciação genética entre as populações;  $R_{st}>0,15$  – alta diferenciação genética entre as populações. (+)  $p<0,05$ .

FONTE: O autor (2019).

A média dos valores de  $R_{st}$  específicos por população, todos superiores ao valor de 0,05, indicam a existência de estrutura populacional, sendo esta mais pronunciada para GKW (0,116) e menos para KIV (0,057) (Tabela 6).

A presença de variabilidade genética associada à estimativa do valor médio de  $R_{st}$  de -0,004 ( $p>0,5$ ), obtido para as 5 populações ameríndias em

conjunto, indica ausência de endogamia nestas populações (Tabela 6). Esta leitura se mostra coerente com os resultados obtidos na investigação do equilíbrio de HW, na qual não se observou excesso de homozigotos para os marcadores utilizados. Desta forma, pode-se afirmar que a diferenciação verificada nestas populações ameríndias se deve, em maior proporção, à deriva genética, como o efeito fundador, característico do povoamento das Américas.

**TABELA 6 - VALORES DE Rst E Ris PARA AS POPULAÇÕES AMERÍNDIAS.**

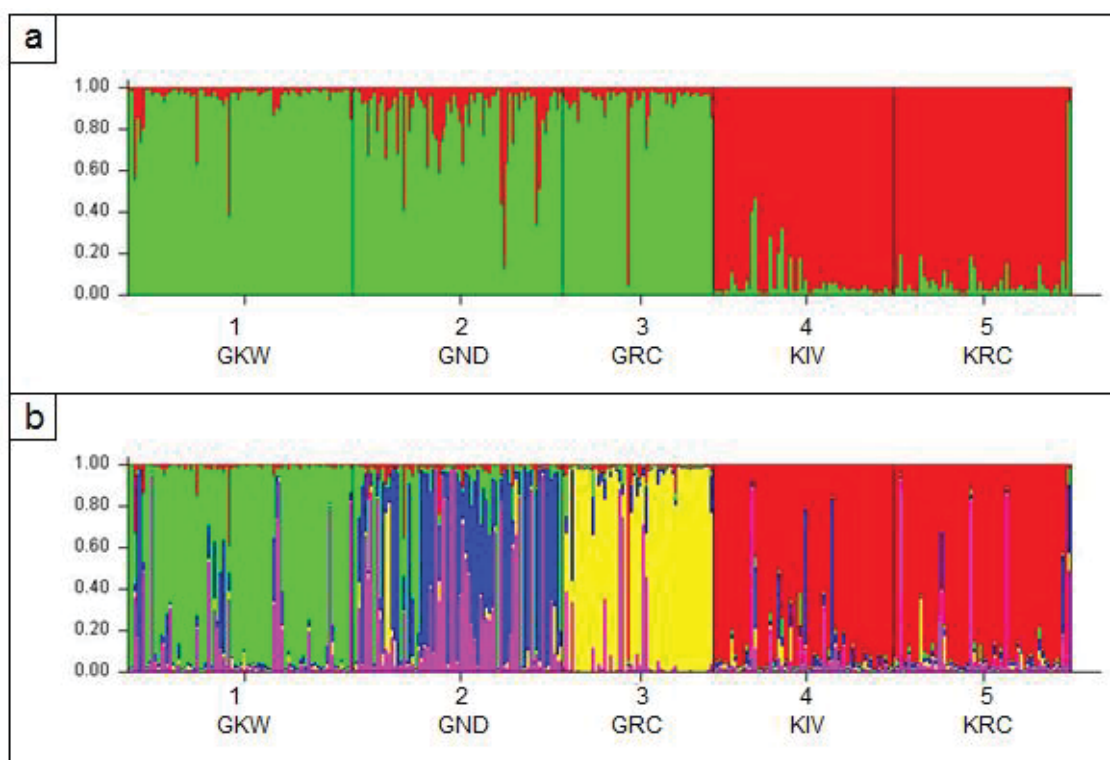
População ameríndia	N	Rst (média individual por população ameríndia)	Ris (média conjunta para as populações ameríndias)
GKW	76	0,116	-0,004
GND	71	0,061	
GRC	51	0,083	
KIV	61	0,057	
KRC	60	0,079	

GKW: Guaraní Kaiowá; GND: Guaraní Ñandeva; GRC: Guaraní M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. N: número de indivíduos.

FONTE: O autor (2019).

### 5.2.2 Structure

A análise de agrupamento realizada no programa STRUCTURE entre os ameríndios, inferiu que o melhor número de agrupamentos (K), é K=2 (Delta K = 240,059233), sendo um dos grupos formado pelos Guaranis e o outro pelos Kaingang. Tal agrupamento revela similaridade genética entre os Guaranis entre si e dos Kaingangs entre si, e evidencia as diferenças entre estes (Figura 5a). Quando os genótipos foram analisados usando um modelo com 5 agrupamentos, fica evidente que os Kaingangs são muito mais parecidos entre si do que os Guaranis entre si (Figura 5b). Dentre as populações Guaraní, GND apresenta maior proporção de compartilhamento genético com Kaingangs. Dentre os Kaingangs, KIV apresenta a maior proporção de compartilhamento genético com Guaranis. Os valores de Delta K, obtidos para os números de agrupamentos, se encontram disponibilizados no Anexo VI (Tabela 26).



**FIGURA 5** - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA ESTIMADO PARA OS INDIVÍDUOS E PARA AS POPULAÇÕES DE AMERÍNDIOS GUARANIS (GKW, GND E GRC) E KAINGANGS (KRC E KIV).

a) Considerando  $K=2$ . Cada indivíduo é representado por uma linha vertical com duas cores (verde e vermelho), em que a proporção de cada cor está relacionada às proporções de compartilhamento genético de cada um dos grupos populacionais inferidos.

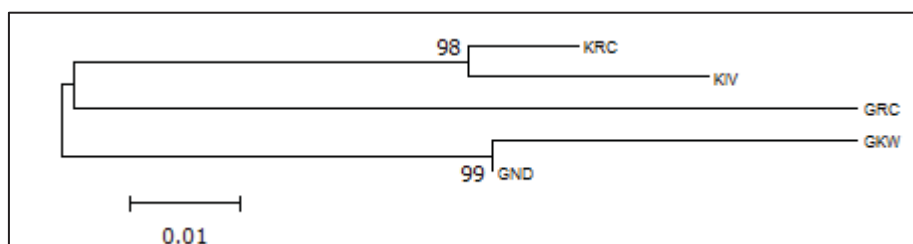
b) Considerando  $K=5$ . Cada indivíduo é representado por uma linha vertical com duas cores (verde, vermelho, azul, amarelo e lilás), em que a proporção de cada cor está relacionada às proporções de compartilhamento genético de cada um dos grupos populacionais inferidos.

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras.

FONTE: O autor (2019).

### 5.2.3 Agrupamento hierárquico

Na análise de agrupamento hierárquico, KIV e KRC aparecem agrupados a partir do mesmo ramo, confirmando sua similaridade e proximidade genética. As populações Guarani são representadas em um agrupamento paralelo a partir do qual GRC é estabelecido como a população mais divergente das três populações Guaranis (Figura 6).



**FIGURA 6** - DENDROGRAMA GERADO COM BASE NOS MARCADORES 21 STR, PARA AS POPULAÇÕES INDÍGENAS.

Foi utilizando o método de agrupamento *Neighbor-Joining* com *bootstrap* de 1000 randomizações.

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras.

FONTE: O autor (2019).

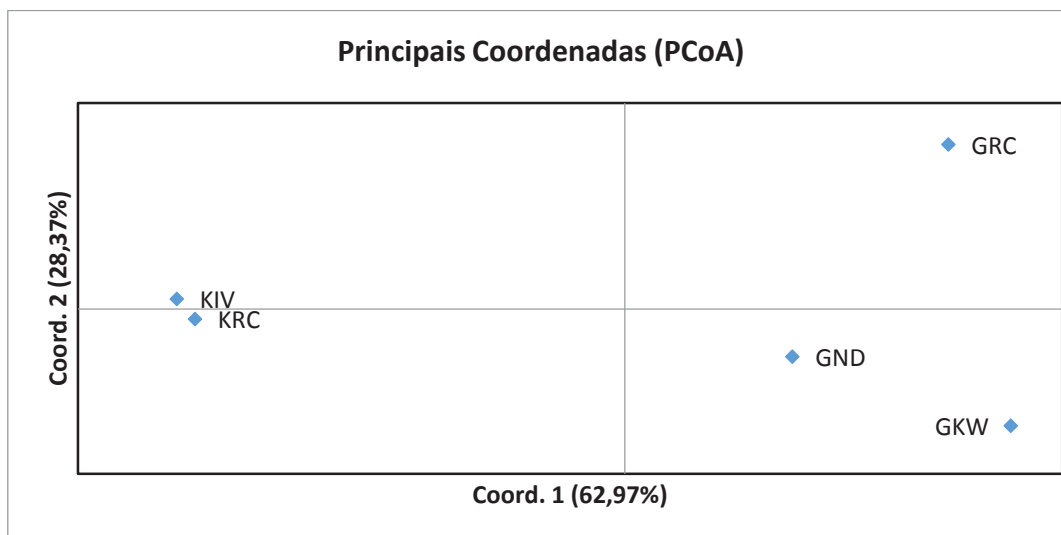
#### 5.2.4 PCoA

Análises de coordenadas principais (PCoA) foram realizadas nas matrizes de genótipos dos grupos populacionais, buscando-se relacionar os resultados dessa análise com os valores de  $R_{ST}$  estimados, uma vez que ambos permitem investigar a estruturação genética.

A PCoA demonstrou que a coordenada 1 (Coord1) explica 62,97% da variação e a coordenada 2 (Coord2) explica 28,37% da variação entre os grupos (Figura 7). Nos grupos Kaingangs e Guaranis, que apresentaram valores altos de  $R_{ST}$ , a Coord1 diferenciou as populações que os compõem.

Guaranis são sempre representados distantes de Kaingangs e, estes são sempre dispostos próximos entre si, representando um único agrupamento. GKW e GND são representados mais próximos entre si do que com relação a GRC (Figura 7).





**FIGURA 7** - PCOA BASEADA EM 21 STRS DEMONSTRANDO PADRÕES DE VARIAÇÃO GENÉTICA NAS POPULAÇÕES AMERÍNDIAS ESTUDADAS.

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras.

FONTE: O autor (2019).

Assim, KIV e KRC podem ser compreendidos como um único conjunto, GND e GKW como outro agrupamento, e GRC como um terceiro grupo. A população GND, na qual se evidenciou maior compartilhamento genético com Kaingangs, através do programa Structure, é representada de forma mais próxima de Kaingangs do que os demais Guaranis, mas ainda assim está mais próxima de GKW. Esta representação é consistente com a constatada pela estatística R, com a observada no dendograma e com os resultados do programa Structure, correspondendo também à dinâmica populacional esperada dados a disposição geográfica e os hábitos culturais de cada população.

GKW e GND estão localizadas em regiões próximas no Mato Grosso do Sul, com distâncias oscilando entre 437 km (Amambai a Limão Verde) e 130 km (Amambai a Porto Lindo), o que poderia facilitar ou mesmo predispor a ocorrência de fluxo gênico entre elas, factível se for considerado o pertencimento de ambas à mesma etnia original. Foi constatada diferenciação entre elas, mas esta foi de baixa magnitude ( $R_{st}=0,04$ ). Este achado poderia indicar, entre GKW e GND, algum nível discreto de fluxo gênico ou, possivelmente, a existência de um ancestral comum mais recente.

A distância geográfica (cerca de 485 km) entre GND (no Mato Grosso do Sul) e GRC (no Paraná), e entre GRC e GKW (Mato Grosso do Sul) não representaria necessariamente um impeditivo a migrações entre estas comunidades. Contudo, foi constatada diferenciação, sendo de média magnitude entre GND e GRC e de alta magnitude ( $R_{st}=0,1557$ ) entre GKW e GRC. A diferenciação verificada demonstra que praticamente não há fluxo gênico entre estas.

Uma vez que a distância física representaria um obstáculo pequeno à ocorrência de fluxo gênico entre as populações Guaranis, supõe-se que as diferenças culturais existentes (INSTITUTO-SOCIOAMBIENTAL, 2016; PETZL-ERLER *et al*, 1993), exerçam pressão maior pelo respectivo isolamento.

A diferenciação muito baixa constatada entre o par das duas populações Kaingangs ( $R_{st}=0,0017$ ), pode ser considerada bastante sugestiva da ocorrência de fluxo gênico, sendo esta possibilidade viabilizada pela proximidade geográfica (ambas se localizam no estado do Paraná, a 300 km de distância) e pelo compartilhamento de características culturais e linguísticas. Adicionalmente, constam relatos de que as famílias Kaingang mudam-se frequentemente de aldeia (INSTITUTO-SOCIOAMBIENTAL, 2016), predispondo-as a esta dinâmica populacional.

Entre as populações Guarani e Kaingang a diferenciação varia de média (entre os pares KIV com relação a cada uma das populações Guarani; e entre KRC e os Guaranis GND e GRC) a alta (entre KRC e GKW). Tal achado também desfavorece a hipótese de fluxo gênico entre Guaranis e Kaingangs e sugere mais intensamente o isolamento entre estas populações nitidamente distintas entre si pelo grupo linguístico a que pertencem. Contudo, entre GND (Mato Grosso do Sul) e Kaingangs (Paraná), resultados obtidos pelo programa Structure sugerem similaridade genética mesmo tendo sido constatada média diferenciação ( $R_{st}=0,07$ ).

Em concordância com estes achados, publicações anteriores, conduzidas para marcadores HLA, relataram a ocorrência de fluxo gênico entre os ameríndios como sendo relativamente pequena (aproximadamente 1,4% dos haplótipos HLA presentes nos Guaranis seriam de origem Kaingang e 0,5% do *pool* gênico dos Kaingangs seria de origem Guarani) (PETZL-ERLER *et al*, 1993).

A existência de diferenciação das três populações Guaranis entre si sugere que se apresentam como populações geneticamente diferentes, embora pertençam à mesma denominação. As populações KIV e KRC, pouco diferenciadas, podem ser consideradas uma única população, no que concerne à diferenciação genética, para os marcadores aqui analisados. Tais resultados confirmam conjecturas anteriormente levantadas a partir de marcadores uniparentais (MARRERO *et al*, 2007).

Portanto, o repetido padrão de similaridade e proximidade genética entre Kaingangs através de todas as análises realizadas permite concluir que, com relação aos marcadores STR forenses, Kaingangs devam ser considerados como uma única população.

Foi então gerada uma nova análise de Rst par a par, desta vez considerando KIV e KRC como uma população única (KAI) (Tabela 7).

**TABELA 7** - VALORES DE DIFERENCIAÇÃO (RST) PARA 4 POPULAÇÕES INDÍGENAS.

	GKW	GND	GRC	KAI
GKW	-	+	+	+
GND	0,04019	-	+	+
GRC	0,15571	0,05544	-	+
KAI	0,13383	0,07345	0,05853	-

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KAI: Kaingangs unificados;  $R_{st} < 0,05$  – baixa diferenciação genética.  $0,05 < R_{st} < 0,15$  - média diferenciação genética entre as populações;  $R_{st} > 0,15$  - alta diferenciação genética entre as populações. (+)  $p < 0,05$ .

FONTE: O autor (2019).

Quando unificados, Kaingangs apresentam média diferenciação para com as três populações Guaranis. O valor mais alto foi constatado para GKW ( $R_{st}=0,13$ ). Com relação a GND e GRC, obteve-se, para KAI, valores similares entre si:  $R_{st}=0,07$  e  $R_{st}=0,058$ , respectivamente.

#### 5.4 PARÂMETROS FORENSES

Os parâmetros forenses foram estimados para cada população indígena individualmente (GKW, GND, GRC, KRC e KIV). Foram também estimados o

número de alelos encontrados em cada população e o respectivo número efetivo de alelos (número de alelos igualmente frequentes que seriam necessários para produzir a heterozigosidade esperada nas populações estudadas) (Tabela 8).

As heterozigosidades médias, mostraram-se semelhantes para as cinco populações ameríndias entre si (Tabela 8), com valor de 0,69 (o valor mais alto) para cada um dos Kaingangs e variando, de 0,67 (GKW, o valor mais alto para Guaranis) a 0,63 (GRC, o valor mais baixo).

Os resultados de heterozigosidades médias, traduzindo a variabilidade genética de cada população, se equiparam aos já relatados para as mesmas populações, com discretas diferenças (Guarani 0,71 e Kaingang 0,77) possivelmente decorrentes de estas cinco populações terem sido simplesmente avaliadas como sendo duas, Guaranis e Kaingangs (Kohlrausch *et al*, 2004; WANG *et al*, 2007).

**TABELA 8 – REGIÃO GEOGRÁFICA, NÚMERO DE ALELOS, HETEROZIGOSIDADE ESPERADA E HETEROZIGOSIDADE OBSERVADA DAS POPULAÇÕES GUARANI E KAINGANG**

Região Geográfica	População Ameríndia	N	Na	HE	HO
Amambai e Limão Verde	GKW	76	6,95	0,68	0,67
Amambai e Porto Lindo	GND	71	7,05	0,70	0,66
Rio das Cobras	GRC	51	5,62	0,65	0,63
Ivaí	KIV	61	6,00	0,69	0,69
Rio das Cobras	KRC	60	6,33	0,68	0,69

GKW: Guarani Kaiowá; GND: Guarani Ñandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. N: número de indivíduos. NA: número de alelos médio por marcador. HE: Heterozigosidade Esperada. HO: Heterozigosidade observada.

Fonte: O autor (2019).

Os marcadores com maior diversidade genética (heterozigose esperada) foram: SE33 (para GKW e KRC), D1S1656 (para GND e KIV) e FGA (para GRC) (Tabela 9).

O marcador com menor heterozigosidade para as três populações Guarani (GKW, GND e GRC) foi D2S441. Já para KIV e KRC, observou-se que os marcadores com menor heterozigosidade foram D22S1045 e D3S1358, respectivamente (Tabela 9).

Nas populações ameríndias GKW e KRC o marcador SE33 foi o mais informativo, bem como se considerados Kaingangs unificados, assim como também relatado para parte dos indígenas norte americanos (Apache, Cora e Yavapai) e para a população brasileira (KANTHASWAMY *et al*, 2016; GUSMÃO *et al*, 2016). Para GND os dois marcadores mais informativos foram SE33 e D1S1656. Para KIV, o marcador SE33 apresentou menor PC e maiores valores para PD e PIC, enquanto D1S1656 apresentou maiores valores para PE e PI (Tabela 9).

Os marcadores menos informativos foram D2S441 (para as populações Guaraní: GKW, GND e GRC), vWA e D22S1045 (para KIV e Kaingangs unificados) e D22S1045 e D3S1358 (para KRC) (Tabela 9). Na população ameríndia mexicana, D2S441 e D22S1045 também foram os marcadores menos informativos (BURGUETE-ARGUETA *et al*, 2015).

**TABELA 9** - MARCADORES FORENSES MAIS E MENOS INFORMATIVOS PARA CADA POPULAÇÃO INDÍGENA.

Parâmetros forenses	GKW	GND	GRC	KIV	KRC	KAI
	<b>Marcadores Mais Informativos</b>					
<b>Menor PC</b>	SE33	SE33	FGA	SE33	SE33	SE33
<b>Maior PD</b>	SE33	SE33	FGA	SE33	SE33	SE33
<b>Maior PIC</b>	SE33	SE33	FGA	SE33	SE33	SE33
<b>Maior PE</b>	SE33	D1S1656	FGA	D1S1656	SE33	SE33
<b>Maior IP</b>	SE33	D1S1656	FGA	D1S1656	SE33	SE33
<b>Maior HE</b>	SE33	D1S1656	FGA	D1S1656	SE33	SE33
	<b>Marcadores Menos Informativos</b>					
<b>Maior PC</b>	D2S441	D2S441	D2S441	vWA	D22S1045	vWA
<b>Menor PD</b>	D2S441	D2S441	D2S441	vWA	D22S1045	vWA
<b>Menor PIC</b>	D2S441	D2S441	D2S441	vWA	D22S1045	D22S1045
<b>Menor PE</b>	D2S441	D2S441	D2S441	D22S1045	D3S1358	D22S1045
<b>Menor IP</b>	D2S441	D2S441	D2S441	D22S1045	D3S1358	D22S1045
<b>Menor HE</b>	D2S441	D2S441	D2S441	D22S1045	D3S1358	D22S1045

GKW: Guaraní Kaiowá; GND: Guaraní Nandeva; GRC: Guaraní M'byá; KIV: Kaingang de Ivai; KRC: Kaingang de Rio das Cobras; KAI: Populações Kaingang unificadas.

FONTE: O autor (2019).

Os valores do poder de discriminação combinado e do poder de exclusão combinado, para os marcadores STR nas populações indígenas,

foram determinados considerando-se Guaranis como três populações distintas entre si e Kaingangs como uma população unificada (Tabela 10).

**TABELA 10** - VALOR DE PODER DE DISCRIMINAÇÃO COMBINADO (PDC) E PODER DE EXCLUSÃO COMBINADO (PEC), ESTIMADOS PARA O CONJUNTO DE 21 MARCADORES STR, POR POPULAÇÃO INDÍGENA.

	GKW	GND	GRC	KAI
<b>PDC</b>	0,9999999999999996	0,9999999999999998	0,9999999999999995	0,9999999999999998
<b>PEC</b>	0,99999664	0,9999849	0,99995705	0,999996447

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras.

FONTE: O autor (2019).

Todos os demais parâmetros forenses encontram-se detalhados no Anexo II (Tabela 21).

## 5.5 DISTRIBUIÇÃO DA DIVERSIDADE GENÉTICA DE GUARANIS E KAINGANGS COM RELAÇÃO A OUTRAS POPULAÇÕES

### 5.5.1 Populações Brasileiras Guarani, Kaingang e Urbana Miscigenada

Foram comparadas as frequências alélicas das populações indígenas com populações brasileiras miscigenadas representantes das cinco regiões geográficas brasileiras (Tabela 11). As frequências alélicas mais elevadas, observadas para a maioria dos marcadores dentre os indígenas, permite inferir que estes apresentam diversidade menor. Tal constatação também foi relatada para marcadores HLA (TSUNETO *et al*, 2003; PETZL-ERLER *et al*, 1993).

Por outro lado, a miscigenação existente na população brasileira, com a contribuição genética de várias origens populacionais, proporciona a esta uma variedade maior no número de alelos observados individualmente por marcador.

TABELA 11 – Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

		(continua)									
Marcador	Alélo	GKW (76)	GND (71)	GRC (51)	KIV (61)	KRC (60)	S (268)	SE (211)	N (182)	NE (241)	CO (240)
D3S1358	9	0	0	0	0	0	0	0,00237	0,00275	0	0
	11	0	0	0	0	0	0	0,00237	0	0	0
	12	0	0	0	0,0082	0	0,00187	0,00237	0	0,01037	0,00208
	13	0	0	0	0	0	0,00373	0,00474	0,00275	0,00415	0,00417
	14	0,05263	0,07746	0	0	0	0,10261	0,04502	0,06319	0,10166	0,07292
	15	0,42763	0,49296	0,33333	0,57377	0,66667	0,29478	0,32464	0,34615	0,28423	0,33333
	16	0,33553	0,23944	0,55882	0,2377	0,275	0,23507	0,27725	0,31319	0,28423	0,30208
	17	0,13816	0,1338	0,03922	0,18033	0,05833	0,21828	0,22275	0,17308	0,17842	0,17083
	18	0,04605	0,0493	0,04902	0	0	0,13246	0,11137	0,08516	0,12656	0,1
	19	0	0,00704	0,01961	0	0	0,00746	0,00237	0,01099	0,0083	0,01458
	20	0	0	0	0	0	0,00373	0,00237	0	0,00207	0
	21	0	0	0	0	0	0	0,00237	0	0	0
	152	0	0	0	0	0	0	0	0,00275	0	0
TH01	5	0	0	0	0	0	0	0	0	0	0,00214
	6	0,60526	0,60563	0,34314	0,11475	0,10833	0,26679	0,19668	0,26685	0,19917	0,18162
	7	0,28289	0,30282	0,29412	0,41803	0,45833	0,23321	0,28199	0,24438	0,25726	0,27564
	8	0,00658	0,01408	0	0	0	0,13619	0,1564	0,16011	0,14315	0,16453
	9	0,00658	0,01408	0	0,0082	0,01667	0,125	0,13507	0,12079	0,19295	0,13462
	10	0	0	0	0	0	0,0056	0,00948	0,01404	0,0083	0,01068
	53	0,00658	0	0	0,0082	0	0	0	0	0	0
	63	0,00658	0	0	0	0	0	0,00474	0	0	0
	83	0	0	0	0	0	0	0	0	0	0,00214
	93	0,08553	0,06338	0,36275	0,45082	0,41667	0,23321	0,21327	0,19382	0,19917	0,22863
	103	0	0	0	0	0	0	0,00237	0	0	0

TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

D21S11	(continuação)														
25	0	0	0	0	0	0	0,00474	0	0,00415	0	0	0,00208	0,00415	0	0
26	0	0	0	0	0	0	0,00237	0	0	0,00208	0,00208	0,00208	0,00208	0,00208	0,00208
27	0	0	0	0	0,05738	0,03333	0,03318	0,01374	0,0249	0,03333	0,03333	0,03333	0,0249	0,03333	0,03333
28	0	0	0	0,0098	0,04918	0,08333	0,17299	0,11264	0,15975	0,17292	0,17292	0,17292	0,15975	0,17292	0,17292
29	0,05263	0,11972	0,26471	0,09016	0,09016	0,05833	0,18246	0,23901	0,18465	0,16667	0,16667	0,16667	0,18465	0,16667	0,16667
30	0,33553	0,47183	0,36275	0,42623	0,42623	0,48333	0,25355	0,25824	0,23859	0,27292	0,27292	0,27292	0,23859	0,27292	0,27292
31	0	0,01408	0,13725	0	0	0,025	0,06398	0,0467	0,07054	0,05833	0,05833	0,05833	0,07054	0,05833	0,05833
32	0	0	0	0	0	0	0,00948	0,00824	0,0083	0,00417	0,00417	0,00417	0,0083	0,00417	0,00417
34	0	0	0	0	0	0	0,00237	0,00549	0,00207	0	0	0	0,00207	0	0
35	0	0	0	0	0	0	0,01185	0,00549	0,01037	0,01042	0,01042	0,01042	0,01037	0,01042	0,01042
36	0	0	0	0,0082	0,0082	0	0	0	0,00622	0	0	0	0,00622	0	0
37	0	0	0	0	0	0	0	0,00275	0	0,00208	0,00208	0,00208	0	0,00208	0,00208
39	0	0	0	0	0	0	0,00237	0	0	0	0	0	0	0	0
242	0	0	0	0	0	0	0,00187	0	0,00207	0	0	0	0,00207	0	0
243	0	0	0	0	0	0	0,00187	0,00275	0,00415	0,00417	0,00417	0,00417	0,00415	0,00417	0,00417
252	0	0	0	0	0	0	0,00187	0	0,00207	0,00208	0,00208	0,00208	0,00207	0,00208	0,00208
283	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0	0
291	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0	0
292	0	0	0	0	0	0	0	0	0	0,00417	0,00417	0,00417	0	0,00417	0,00417
293	0	0	0	0	0	0	0	0	0,00207	0	0	0	0,00207	0	0
302	0,00658	0,00704	0	0	0	0	0,03358	0,01923	0,03734	0,02292	0,02292	0,02292	0,03734	0,02292	0,02292
312	0,05921	0,07746	0,02941	0,04918	0,04918	0,025	0,09328	0,1044	0,12448	0,11458	0,11458	0,11458	0,12448	0,11458	0,11458
322	0,28289	0,17606	0,16667	0,14754	0,14754	0,09167	0,09701	0,13187	0,07884	0,08958	0,08958	0,08958	0,07884	0,08958	0,08958
331	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0	0
332	0,25658	0,1338	0,02941	0,17213	0,17213	0,175	0,03172	0,03846	0,03112	0,03542	0,03542	0,03542	0,03112	0,03542	0,03542
342	0	0	0	0	0	0	0,00373	0,01099	0,00622	0,00417	0,00417	0,00417	0,00622	0,00417	0,00417











TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

	(continuação)														
	102	0	0	0	0	0	0,00188	0	0	0	0,00188	0,00237	0	0	0
112	0	0	0	0	0	0	0,00188	0,00237	0	0	0,00188	0,00237	0	0,00208	0,00216
122	0	0	0	0	0	0	0,01128	0,00237	0,01143	0	0,01128	0,00237	0,01143	0,00417	0,00433
132	0	0	0	0	0	0	0,00564	0,00237	0	0	0,00564	0,00237	0	0,00208	0
142	0	0	0	0	0	0	0,00188	0,00237	0	0	0,00188	0,00237	0	0,00208	0,00216
152	0	0	0	0	0	0	0,00188	0,00237	0,00571	0	0,00188	0,00237	0,00571	0,00208	0,00216
162	0	0	0	0	0	0	0	0,00474	0,00286	0	0	0,00474	0,00286	0	0,00216
163	0	0	0	0	0	0	0,00188	0	0	0	0,00188	0	0	0,00208	0,00216
173	0	0	0	0	0	0	0	0,00237	0	0	0	0,00237	0	0	0,00216
181	0	0	0	0	0	0	0	0,00237	0	0	0	0,00237	0	0	0
182	0	0	0	0	0	0	0,00188	0	0	0	0,00188	0	0	0,00208	0
183	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00208	0
192	0	0	0	0	0	0	0,00188	0	0	0	0,00188	0	0	0,00208	0,00216
193	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00208	0
202	0	0	0	0	0	0	0,01128	0,00474	0,00286	0	0,01128	0,00474	0,00286	0	0,00866
212	0	0	0	0	0	0	0,01128	0,02133	0,00286	0	0,01128	0,02133	0,00286	0,00833	0,00649
222	0	0	0	0	0	0	0,04323	0,01896	0,02	0	0,04323	0,01896	0,02	0,01667	0,01732
232	0	0	0	0	0	0	0,03008	0,02607	0,01429	0	0,03008	0,02607	0,01429	0,01458	0,03247
242	0	0	0	0	0	0	0,00833	0,03791	0,02857	0,00833	0,03008	0,03791	0,02857	0,025	0,02381
252	0,02632	0	0	0	0,09167	0	0,05833	0,02133	0,02857	0	0,04323	0,02133	0,02857	0,03125	0,0368
262	0	0,08451	0	0,03333	0,03333	0	0,04887	0,08294	0,05714	0,03333	0,04887	0,08294	0,05714	0,05833	0,0303
272	0,07895	0,05634	0,02941	0,02941	0,225	0,05833	0,05639	0,05687	0,04571	0,225	0,05639	0,05687	0,04571	0,07292	0,05628
282	0,06579	0,11268	0,2451	0,16667	0,18333	0,18333	0,05827	0,05687	0,07429	0,16667	0,05827	0,05687	0,07429	0,08333	0,06277
292	0,01974	0	0,0098	0,025	0	0	0,05263	0,04502	0,03714	0,025	0,05263	0,04502	0,03714	0,03958	0,06277
302	0,01316	0,02817	0,15686	0,025	0,025	0,14167	0,03947	0,02607	0,04	0,025	0,03947	0,02607	0,04	0,04167	0,0303
312	0,04605	0	0	0	0	0	0,03008	0,02844	0,01429	0	0,03008	0,02844	0,01429	0,01667	0,02814

**TABELA 11** - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

	(continuação)										
vWA	322	0,00658	0,02113	0	0	0	0,01504	0,01659	0,00857	0,01875	0,01515
	332	0,01316	0,00704	0	0	0	0	0,00474	0,00571	0,00417	0,00866
	342	0	0	0	0	0	0,00188	0,00237	0	0	0,00216
	11	0	0	0	0	0	0,00187	0,00237	0	0,00415	0,00417
	12	0	0	0	0	0	0	0	0,00275	0	0
	13	0	0	0	0	0	0,00373	0,00474	0	0,00415	0
	14	0	0,00704	0,01961	0	0,025	0,10075	0,07346	0,04121	0,07054	0,08333
	15	0	0	0,18627	0	0	0,11567	0,16114	0,12088	0,13485	0,13958
	16	0,28289	0,38028	0,17647	0,53279	0,56667	0,25746	0,23934	0,33791	0,28838	0,27708
	17	0,34211	0,43662	0,5	0,45082	0,39167	0,27612	0,27725	0,32967	0,25311	0,24167
D8S1179	18	0,30263	0,1338	0,11765	0,0082	0,01667	0,15112	0,14218	0,10989	0,17842	0,18125
	19	0,06579	0,04225	0	0,0082	0	0,07836	0,07346	0,04396	0,04979	0,05208
	20	0,00658	0	0	0	0	0,01493	0,0237	0,01099	0,01452	0,01667
	21	0	0	0	0	0	0	0,00237	0,00275	0,00207	0,00417
	8	0	0,01408	0	0	0	0,0056	0,00711	0	0,00415	0,01042
	9	0	0	0	0	0	0,01119	0,00711	0,00824	0,00622	0,00625
	10	0,13816	0,04225	0,06863	0,05738	0,05833	0,08022	0,06872	0,08791	0,06017	0,0875
	11	0	0	0,12745	0	0,03333	0,06716	0,05213	0,04945	0,06017	0,08125
	12	0,04605	0,07042	0,0098	0,21311	0,06667	0,14552	0,16114	0,14011	0,13693	0,14375
	13	0,11184	0,23944	0,26471	0,42623	0,38333	0,30224	0,28673	0,29121	0,26556	0,2625
TPOX	14	0,48684	0,40141	0,48039	0,22131	0,30833	0,23321	0,20616	0,25824	0,24896	0,2375
	15	0,21711	0,21831	0,04902	0,08197	0,15	0,12687	0,16588	0,14286	0,15975	0,13542
	16	0	0,01408	0	0	0	0,02799	0,04028	0,01923	0,05394	0,02708
	17	0	0	0	0	0	0	0,00474	0	0,00415	0,00833
	18	0	0	0	0	0	0	0	0,00275	0	0
	5	0	0	0	0	0,00833	0,00373	0	0	0	0



TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

(continuação)													
	222	0	0	0	0	0	0,00373	0,00474	0,00279	0,00622	0,00426		
	223	0	0	0	0	0	0,00187	0	0	0	0		
	232	0	0	0	0	0	0	0,00237	0	0	0		
	242	0	0	0	0	0	0	0	0,00279	0	0		
	243	0	0	0	0	0	0	0,00237	0	0	0,00213		
	252	0	0	0	0	0	0	0,00237	0	0	0		
	312	0	0	0	0	0	0	0,00237	0,00559	0	0		
	322	0	0	0	0	0	0	0	0	0,00207	0,00213		
	432	0	0	0	0,01639	0	0	0	0,00279	0	0		
	442	0	0	0	0	0	0	0	0	0	0,00213		
	462	0	0	0	0	0	0,00187	0	0	0	0		
	472	0	0	0	0	0	0	0	0	0,00207	0		
D2S1338	14	0	0	0	0	0	0	0,00237	0	0,00415	0,00216		
	15	0	0	0	0	0	0	0	0	0	0,00216		
	16	0,00658	0	0	0	0	0,04104	0,05213	0,03409	0,05187	0,0431		
	17	0,28947	0,32394	0,35294	0,25833	0,225	0,22948	0,22038	0,21023	0,21162	0,20259		
	18	0,04605	0,07746	0,15686	0,05833	0,05	0,08209	0,06635	0,08239	0,07261	0,07974		
	19	0,11842	0,09859	0,07843	0,15833	0,29167	0,11381	0,11848	0,13636	0,139	0,15302		
	20	0,15789	0,11972	0,16667	0,14167	0,08333	0,14739	0,1327	0,15057	0,12448	0,13147		
	21	0	0,01408	0	0	0,01667	0,03731	0,06635	0,05114	0,06017	0,03879		
	22	0,13158	0,19014	0	0,11667	0,09167	0,07836	0,0782	0,07386	0,07469	0,07112		
	23	0,22368	0,15493	0,2451	0,26667	0,24167	0,10821	0,09716	0,14205	0,10166	0,14009		
	24	0,01316	0,00704	0	0	0	0,07649	0,07583	0,05398	0,06846	0,0819		
	25	0,01316	0,01408	0	0	0	0,06903	0,07109	0,04545	0,06432	0,0431		
	26	0	0	0	0	0	0,01679	0,01896	0,01705	0,02282	0,00862		
	27	0	0	0	0	0	0	0	0,00284	0,00415	0		

D2S1338



TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

		(continuação)																			
D19S433	242	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00216		
	10	0	0	0	0	0	0	0	0	0	0	0	0	0,00275	0,00415	0,00208	0,00415	0,00208	0,00415	0,00208	
	11	0,00658	0	0	0	0	0	0,00933	0,01422	0,01422	0,11374	0,06319	0,00824	0,01245	0,02292	0,01245	0,02292	0,01245	0,02292	0,01245	0,02292
	12	0	0,02113	0,0098	0,0082	0,00833	0,00833	0,10261	0,25746	0,26066	0,11374	0,06319	0,00824	0,01245	0,02292	0,01245	0,02292	0,01245	0,02292	0,01245	0,02292
	13	0,39474	0,20423	0,66667	0,16393	0,125	0,125	0,25746	0,29478	0,29384	0,26066	0,25824	0,25824	0,25519	0,29167	0,25519	0,29167	0,25519	0,29167	0,25519	0,29167
	14	0,17763	0,1831	0,20588	0,03279	0,03333	0,03333	0,16231	0,29478	0,29384	0,13033	0,16758	0,16758	0,12656	0,15833	0,12656	0,15833	0,12656	0,15833	0,12656	0,15833
	15	0,14474	0,20423	0,05882	0,2459	0,33333	0,33333	0,03172	0,0056	0,00237	0,01185	0,01923	0,01923	0,02905	0,03125	0,02905	0,03125	0,02905	0,03125	0,02905	0,03125
	16	0,01316	0,08451	0,01961	0	0	0	0,0056	0,00237	0,00207	0,00207	0,00207	0,00207	0,00207	0,00208	0,00207	0,00208	0,00207	0,00208	0,00207	0,00208
	17	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0,00208	0	0,00208	0	0,00208	0	0,00208
	52	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0,00208	0	0,00208	0	0,00208	0	0,00208
	102	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00208	0	0,00208	0	0,00208	0	0,00208
	112	0	0	0	0	0	0	0	0	0	0	0	0	0,00275	0,00415	0	0,00275	0,00415	0	0,00275	0,00415
	121	0	0	0	0	0	0	0,00187	0	0	0	0	0	0	0	0	0	0	0	0	0
	122	0,00658	0,05634	0,01961	0,07377	0,03333	0,03333	0,01119	0,0159	0,0159	0,0159	0,02198	0,02198	0,01245	0,01458	0,01245	0,01458	0,01245	0,01458	0,01245	0,01458
	132	0,07237	0,10563	0,01961	0,36885	0,30833	0,30833	0,03731	0,03081	0,03081	0,03081	0,09066	0,09066	0,05602	0,04375	0,05602	0,04375	0,05602	0,04375	0,05602	0,04375
	142	0,01316	0	0	0	0,01667	0,01667	0,01866	0,05213	0,05213	0,05213	0,04121	0,04121	0,04772	0,02292	0,04772	0,02292	0,04772	0,02292	0,04772	0,02292
152	0,17105	0,1338	0	0,10656	0,14167	0,14167	0,05037	0,04976	0,04976	0,04976	0,0522	0,0522	0,07676	0,05417	0,07676	0,05417	0,07676	0,05417	0,07676	0,05417	
162	0	0,00704	0	0	0	0	0,01493	0,0237	0,0237	0,0237	0,00824	0,00824	0,0166	0,01667	0,0166	0,01667	0,0166	0,01667	0,0166	0,01667	
172	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00417	0	0,00417	0	0,00417	0	0,00417	
D10S1248	7	0	0	0	0	0,00833	0,00833	0	0	0	0,00237	0	0,00549	0	0	0	0	0	0	0	
	8	0	0	0	0	0	0	0	0	0,00237	0	0	0	0	0	0	0	0	0	0	
	9	0	0	0	0	0	0	0,00187	0,00237	0,00237	0,00237	0,00275	0,00275	0,00207	0,00208	0,00207	0,00208	0,00207	0,00208	0,00207	0,00208
	10	0	0	0	0	0	0	0	0,00237	0,00237	0,00237	0,00275	0,00275	0,00415	0	0,00415	0	0	0,00415	0	
	11	0	0	0	0	0,025	0,025	0,00373	0,0237	0,0237	0,0237	0,01374	0,01374	0,03527	0,01875	0,03527	0,01875	0,03527	0,01875	0,03527	0,01875
	12	0	0,02113	0	0,0082	0,00833	0,00833	0,04664	0,06635	0,06635	0,06635	0,03297	0,03297	0,05394	0,06458	0,05394	0,06458	0,05394	0,06458	0,05394	0,06458
	13	0,38816	0,11972	0,29412	0,04098	0,08333	0,08333	0,26679	0,27725	0,27725	0,27725	0,27747	0,27747	0,28423	0,25833	0,28423	0,25833	0,28423	0,25833	0,28423	0,25833



TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.

(continuação)											
D2S441	143	0	0	0	0	0	0,00373	0,00711	0,00824	0,01037	0,00208
	153	0	0	0	0	0,00833	0,04664	0,04739	0,02747	0,05394	0,03958
	163	0,11842	0,11268	0	0	0	0,04478	0,09242	0,0467	0,05394	0,08125
	171	0	0	0	0	0	0	0	0	0,00207	0
	173	0,26974	0,20423	0,26471	0,35246	0,35	0,14925	0,1019	0,18956	0,12448	0,13958
	183	0,03947	0,05634	0,13725	0,17213	0,10833	0,04478	0,04502	0,04945	0,03734	0,04583
	193	0	0	0	0	0	0,00933	0,01185	0,00824	0,01245	0,00417
	203	0	0	0	0	0	0,00187	0	0,00549	0	0,00417
	8	0	0	0	0	0	0	0	0,00275	0	0
	9	0	0	0	0	0	0	0,00474	0	0,00415	0,00208
	10	0,78947	0,73944	0,87255	0,38525	0,38333	0,25933	0,22512	0,40385	0,21992	0,23958
	11	0,125	0,16197	0,07843	0,56557	0,53333	0,32836	0,30332	0,23901	0,32365	0,30208
D12S391	12	0,07237	0,02817	0	0,0082	0,025	0,04478	0,06872	0,03846	0,06017	0,07708
	13	0,00658	0,00704	0	0	0	0,02239	0,03081	0,02198	0,02697	0,03333
	14	0,00658	0,05634	0,04902	0,04098	0,05833	0,23507	0,26066	0,23352	0,28216	0,25
	15	0	0	0	0	0	0,04104	0,04976	0,02473	0,0249	0,02292
	16	0	0	0	0	0	0,00373	0,00237	0	0,00415	0,00208
	113	0	0,00704	0	0	0	0,0597	0,05213	0,03297	0,04979	0,06042
	123	0	0	0	0	0	0,0056	0,00237	0,00275	0,00207	0,01042
	133	0	0	0	0	0	0	0	0	0,00207	0
	15	0	0	0	0	0	0,03545	0,03791	0,01944	0,06017	0,04603
	16	0,00658	0	0	0,05833	0,1	0,03731	0,03318	0,03889	0,04149	0,03347
	17	0	0,01408	0	0,00833	0	0,11567	0,09953	0,08889	0,09959	0,1046
	18	0,05263	0,07042	0,06	0,05833	0,025	0,18097	0,18957	0,20833	0,17635	0,19665
19	0,21053	0,42254	0,39	0,45833	0,275	0,15672	0,13981	0,22222	0,13693	0,159	
20	0,42763	0,37324	0,24	0,28333	0,23333	0,11567	0,14929	0,16389	0,1639	0,15481	

D12S391

D2S441

**TABELA 11 - Frequências alélicas das populações Guarani e Kaingang e brasileira miscigenada, para os 21 marcadores STR.**

																	(conclusão)
21	0,06579	0,06338	0,08	0,11667	0,21667	0,07836	0,1019	0,075	0,09959	0,09623							
22	0,02632	0,02817	0,23	0	0,03333	0,10821	0,08531	0,06389	0,10788	0,06276							
23	0,17763	0,02113	0	0,01667	0,09167	0,05224	0,06398	0,05	0,05809	0,06695							
24	0,02632	0,00704	0	0	0	0,02799	0,01896	0,01667	0,01037	0,01674							
25	0	0	0	0	0	0,02612	0,01185	0,01111	0,00415	0,01046							
26	0	0	0	0	0	0,00746	0,00237	0,00556	0	0							
27	0	0	0	0	0	0,00187	0	0	0	0							
151	0	0	0	0	0	0	0,00237	0	0	0,00209							
163	0	0	0	0	0	0	0,00237	0	0	0							
171	0	0	0	0	0	0,00373	0	0,00278	0,00415	0							
173	0	0	0	0	0	0,02052	0,01422	0,00833	0,0083	0,02092							
183	0	0	0	0	0,025	0,01679	0,03081	0,01111	0,01867	0,01464							
191	0	0	0	0	0	0,00187	0,00474	0	0,00207	0,00209							
193	0	0	0	0	0	0,00933	0,00948	0,01389	0,00622	0,01046							
203	0,00658	0	0	0	0	0,00373	0,00237	0	0	0,00209							
213	0	0	0	0	0	0	0	0	0,00207	0							

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras; S: Região Sul; SE: Região Sudeste; N: Região Norte; NE: Região Nordeste; CO: Região Centro-Oeste. Em vermelho: alelo mais frequente para cada marcador, por população.

Os alelos mais frequentes para toda a população indígena, D21S11\*30 e D10S1248\*14, também são os mais frequentes na população brasileira.

Dos alelos mais frequentes nas três populações Guaranis, D2S1338\*17, D19S433\*13 e D2S441\*10, o alelo D2S1338\*17, é também o mais frequente para todas as regiões brasileiras. Nas regiões Nordeste e Centro-Oeste, o alelo mais frequente também foi D19S433\*13 e, na região Norte, igualmente o alelo D2S441\*10.

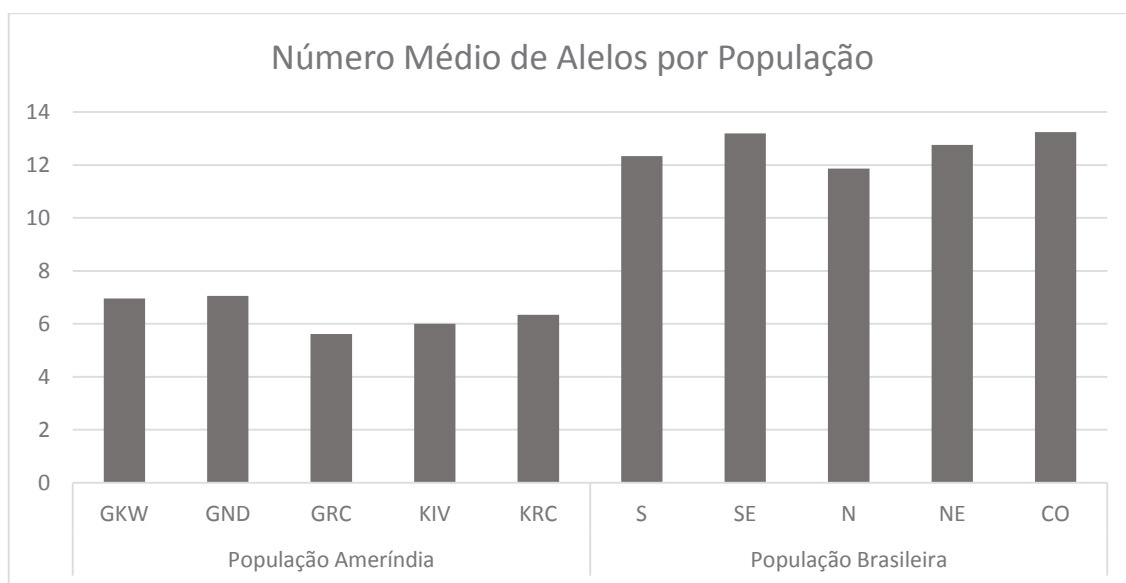
Todas as regiões brasileiras compartilham com Kaingang os alelos D8S1179\*14 e TPOX\*8 como mais frequentes. O alelo D2S441\*11 foi mais frequente entre Kaingangs e todas as regiões, com exceção do Norte. O alelo vWA\*17 foi o mais frequente para Kaingangs e regiões Norte, Nordeste e Centro-Oeste.

Reproduzindo o resultado obtido para as populações ameríndias, o marcador com menor número de alelos na população miscigenada brasileira foi CSF1PO, enquanto o marcador SE33 também apresentou o maior número de alelos. O número médio de alelos, quando consideradas todas as populações brasileiras miscigenadas, foi de 39,6 (CSF1PO) e 9,2 alelos (SE33) (Figura 8).

Entre indígenas e a população miscigenada brasileira, o alelo D22S1045\*11, mostrou-se 20 vezes mais frequente na população nordeste brasileira do que nos GKW, GND e KIV, não tendo sido sequer detectado nos GRC e KRC. O alelo D19S433\*12 é 14 vezes mais frequente na população Nordeste do que entre os Kaingangs e 6 vezes mais frequente do que entre GND e GRC. O alelo TH01\*9 apresenta frequência 31 vezes maior no Nordeste do que em GKW, 13 vezes maior do que para GRC, 23 vezes maior do que para KIV e 11 vezes maior do que em KRC. O alelo SE33\*28 somente foi detectado em duas populações: nos GRC e na população Sudeste, onde é 50 vezes menos frequente.

Foram portanto verificadas desproporções entre as frequências de determinados alelos, quando comparadas as populações indígenas e brasileira miscigenada. São poucos, contudo, estes alelos que se apresentam em proporções tão desproporcionais entre indígenas e a população urbana, sugerindo que, se houve fluxo gênico entre populações urbanas e ameríndias Guarani e Kaingang, este teria sido de relativamente baixa intensidade. A

ocorrência de fluxo gênico de pequena magnitude entre as populações ameríndias e as populações urbanas já havia de fato sido relatado, com base em marcadores do sistema HLA, como sendo de 5,3% para Guarani e 2,5% para Kaingangs (PETZL-ERLER *et al*, 1993). As frequências alélicas observadas nas populações indígenas apresentam-se, para a maioria dos marcadores, maiores do que as observadas na população brasileira. Por outro lado, as populações brasileiras apresentam, com mais frequência, o maior número total e médio de alelos, por marcador (Figura 8).



**FIGURA 8 - NÚMERO MÉDIO DE ALELOS POR POPULAÇÃO AMERÍNDIA E BRASILEIRA MISCIGENADA, CONSIDERANDO O PAINEL DE 21 MARCADORES STR.**

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras; S: População brasileira da região sul; SE: População brasileira da região sudeste; CO: População brasileira da região centro oeste; NE: População brasileira da região nordeste; N: População brasileira da região norte.

FONTE: O autor (2019).

Considerando-se que, numa situação de equilíbrio entre mutação e deriva, a média de número de alelos detectados por marcador equivale a um indicador de variabilidade genética dentro de uma população, comparou-se o número de alelos detectados em cada loco e para cada população (KOHLRAUSCH *et al*, 2005). Através da comparação das frequências alélicas e genotípicas obtidas para as populações indígenas com a população brasileira foi possível constatar que o número observado de alelos em todas as populações ameríndias avaliadas é inferior ao observado para populações

urbanas. Esta constatação está de acordo com o observado para alelos do sistema HLA nestas populações (OLIVEIRA *et al*, 2007). Populações urbanas, particularmente a miscigenada brasileira, apresentam maior diversidade, evidenciada pelo número de alelos que apresenta, por traduzirem uma mistura de diferentes subpopulações (OLIVEIRA *et al*, 2007).

A média total de alelos detectados para cada população Guarani e Kaingang se assemelha à relatada para estes e outros povos ameríndios da América do Sul (KOHLEAUSCH *et al*, 2005; HUTZ *et al*, 2002).

O marcador com menor número de alelos nas populações ameríndias e miscigenada brasileira foi CSF1PO, com média de 4,2 e 9,2 alelos, respectivamente. O marcador SE33 apresentou o maior número de alelos, com média de 13 e 39,6 alelos, respectivamente (Anexo III - Tabela 22). Estes resultados estão de acordo com outras publicações similares, que identificaram CSF1PO como o menos e SE33 como o marcador mais variável.

Comparando-se a distribuição das frequências dos alelos exclusivos detectados entre as populações ameríndias e a população miscigenada brasileira observou-se que, com exceção de um único alelo (D16S539\*15, presente apenas nos GRC), todos se fazem presentes também nas populações urbanas brasileiras. Estas, por sua vez, apresentaram alelos exclusivos que não foram detectados entre os indígenas, fato talvez justificável pela expressiva miscigenação ou pelo descrito isolamento das populações indígenas, o que lhes confere menor variabilidade (PETZL-ERLER *et al*, 1993). De maneira geral, a população miscigenada brasileira, como uma população urbana, apresenta maior proporção de alelos exclusivos para todos os STR testados, em concordância com Wang *et al* (2007). Este fato pode ser compreendido como um contraponto à possibilidade de fluxo gênico entre indígenas e população urbana.

Alelos exclusivos em uma população são detectados quando se comparam várias populações e, neste contexto, tornam-se úteis para o esclarecimento da história demográfica de populações, bem como para estudos de genética voltada à investigação da presença de forças evolutivas, como migração (SZPIECH, 2011). Considerando ainda uma possível associação com a genética forense, alelos exclusivos plenamente estabelecidos e descritos em estudos populacionais poderiam ser úteis como

instrumentos adicionais de presunção da origem populacional de um indivíduo ou de um perfil genético obtido de um vestígio de crime. A informação presumida da procedência populacional de um perfil genético acrescentaria dados importantes às investigações forenses.

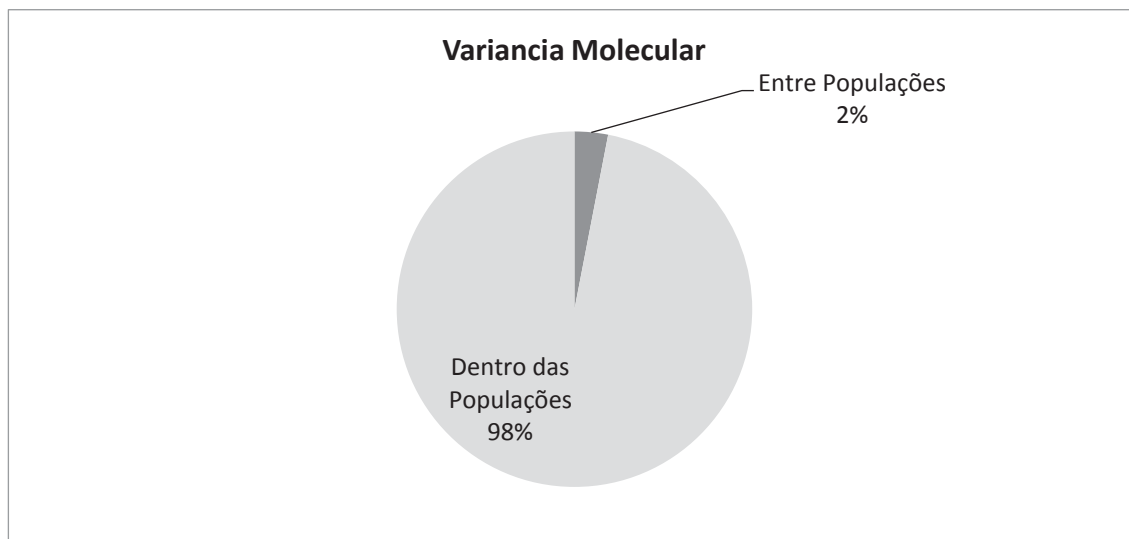
O alelo exclusivo remanescente (D16S539\*15) da comparação entre indígenas e a população miscigenada brasileira, apresenta frequência relativamente baixa (0,029). De todas as populações ameríndias consultadas, somente Conchagua, (El Salvador), apresentou a ocorrência deste mesmo alelo (frequência de 0,0102) (LOVO-GÓMEZ *et al*, 2007). Em diversas populações urbanas europeias este alelo também se encontra reportado, em menor frequência, variando de 0,0048 (Polônia) a 0,0012 (Portugal), assim como no Japão (0,001) e na população siberiana (0,0018) (STRIDER, 2019; TIE *et al*, 2006; ZHIVOTOVSKY *et al*, 2009). A detecção deste alelo em populações situadas na América Central e na Sibéria depõe a favor da hipótese de ondas migratórias originadas na Sibéria e colonizando o continente americano a partir de rotas costeiras de norte a sul (WANG *et al*, 2007). A ocorrência deste alelo em populações na Europa sugere também a possibilidade de convergência. Esta situação difere de outras publicações que permitem deduções relativas à migração, por relatarem a detecção de um alelo específico muito comum às populações ameríndias e ausente nas populações europeias (WANG *et al*, 2007). Não tendo sido encontradas outras coincidências de alelos exclusivos exclusivamente nos ameríndios do presente estudo, considerou-se este dado de forma insuficiente para permitir inferências concretas a respeito da história do povoamento das Américas.

#### 5.5.1.1 AMOVA e Estatística R

Entre a população indígena total e a população miscigenada brasileira foi possível constatar que a maior contribuição à variância total (83% a 100%) se encontra uniformemente distribuída (98%) dentro das populações (Figura 9), para todos os marcadores e em todas as comparações (Tabela 12). Após a correção de Bonferroni ( $P < 0,0024$ ), valores de  $R_{st}$  estatisticamente



significativos foram encontrados para 14 marcadores (TH01, D21S11, D5S818, D7S820, D16S539, SE33, vWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 e D12S391), cuja diferenciação genética ( $R_{st}$  par a par) foi individualmente estimada (Anexo IV – Tabela 24).



**FIGURA 9** - COMPONENTES DE VARIAÇÃO ENTRE POPULAÇÕES E DENTRO DAS POPULAÇÕES, CONSIDERANDO-SE OS INDÍGENAS GUARANI E KAINGANG E A POPULAÇÃO MISCIGENADA BRASILEIRA.

FONTE: O autor (2019).

**TABELA 12** - AMOVA LOCO POR LOCO BASEADO NOS VALORES DE RST CONSIDERANDO AS POPULAÇÕES INDÍGENAS E A POPULAÇÃO MISCIGENADA BRASILEIRA REUNIDAS.

Marcador	Va	% Variação interpopulacional	Vb	% Variação intrapopulacional	Rst	Valor de p
<b>D3S1358</b>	0,000	0%	7,715	100%	-0,024	1,000
<b>TH01</b>	25,936	2%	1236,542	98%	0,021	0,001*
<b>D21S11</b>	368,477	2%	17260,554	98%	0,021	0,001*
<b>D18S51</b>	0,083	0%	60,951	100%	0,001	0,007
<b>D5S818</b>	0,280	11%	2,271	89%	0,110	0,001*
<b>D13S317</b>	0,000	0%	6,374	100%	-0,007	0,016
<b>D7S820</b>	0,000	0%	7,522	100%	-0,009	0,002*
<b>D16S539</b>	0,081	4%	2,095	96%	0,037	0,001*
<b>CSF1PO</b>	0,000	0%	5,606	100%	-0,012	0,007
<b>SE33</b>	197,541	1%	15779,101	99%	0,012	0,001*
<b>VWA</b>	0,013	1%	1,717	99%	0,008	0,001*
<b>D8S1179</b>	0,000	0%	2,379	100%	-0,006	0,001*
<b>TPOX</b>	0,096	3%	3,700	97%	0,025	0,001*
<b>FGA</b>	0,000	0%	885,165	100%	-0,003	0,600
<b>D2S1338</b>	0,000	0%	28,336	100%	-0,005	0,533
<b>D19S433</b>	111,475	4%	2564,595	96%	0,042	0,001*
<b>D10S1248</b>	0,005	0%	1,501	100%	0,003	0,006
<b>D22S1045</b>	0,000	0%	2,465	100%	-0,008	0,001*
<b>D1S1656</b>	36,757	1%	5437,571	99%	0,007	0,001*
<b>D2S441</b>	3,914	1%	448,040	99%	0,009	0,001*
<b>D12S391</b>	3,177	0%	1046,054	100%	0,003	0,002*

Va=variância entre populações; Vb=variâncias individuais dentro da população; \*p<0,0024.

FONTE: O autor (2019).

Entre ameríndios Guarani e a população miscigenada brasileira, constatou-se a existência de diferenciação (Rst) entre todos os pares de populações, com magnitudes variando de baixa ( $0,0043 < Rst < 0,0114$ ) a média ( $0,051 < Rst < 0,1374$ ) (Tabela 13).

Considerados como uma população única, Kaingangs mostram média diferenciação para com todas as regiões brasileiras (Tabela 13).

**TABELA 13** - VALORES DE RST PARA 4 SUBPOPULAÇÕES INDÍGENAS E POPULAÇÃO BRASILEIRA.

	GKW	GND	GRC	KAI	S	SE	N	NE	CO
GKW	-	+	+	+	+	+	+	+	+
GND	0,04019	-	+	+	+	+	+	+	+
GRC	0,15571	0,05544	-	+	-	-	+	+	-
KAI	0,13383	0,07345	0,05853	-	+	+	+	+	+
S	0,13741	0,04084	0,00426	0,05596	-	-	+	-	-
SE	0,13567	0,03662	0,00586	0,05638	0,0005	-	-	-	-
N	0,09401	0,01143	0,01392	0,05723	0,0063	0,0069	-	-	-
NE	0,11563	0,02516	0,01113	0,05134	0,0014	-0,0008	0,0031	-	-
CO	0,12301	0,02925	0,00526	0,05811	0,0007	-0,0008	0,0027	-0,0009	-

GKW: Guarani Kaiowá; GND: Guarani Ñandeva; GRC: Guarani M'byá; KAI: Kaingangs unificados; S: população brasileira da região sul; SE: população brasileira da região sudeste; CO: população brasileira da região centro oeste; NE: população brasileira da região nordeste; N: População brasileira da região norte.  $R_{st} < 0,05$  – baixa diferenciação genética.  $0,05 < R_{st} < 0,15$  – média diferenciação genética entre as populações;  $R_{st} > 0,15$  – alta diferenciação genética entre as populações. (+)  $p < 0,05$ .

FONTE: O autor (2019).

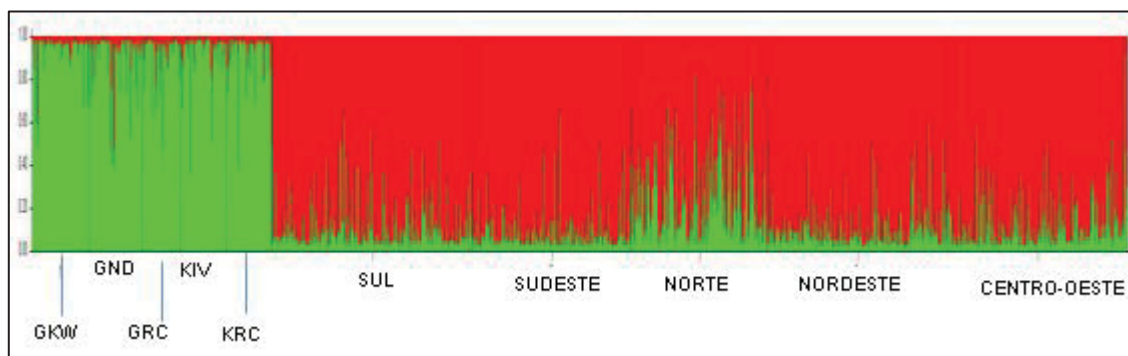
Os resultados de estatística R demonstram que a população miscigenada brasileira apresenta diferenciação para com as populações Guarani e Kaingang, sendo esta diferenciação considerada média (para com GKW e KAI) e baixa (para com GND e GRC), como também relatado por Marrero *et al* (2007). GKW e GND apresentam maior diferenciação para com a região Sul e, menor, para a região Norte. Inversamente, GRC apresenta a maior diferenciação para com a região Norte e, a menor, para a região Sul, demonstrando apresentar características demográficas diferenciadas dos demais Guaranis.

Estes resultados corroboram o observado e relatado para alelos de *DRB1* do sistema HLA: GKW apresenta a maior diferenciação e a menor taxa de mistura com populações não ameríndias; GND apresenta baixa diferenciação e a maior taxa de mistura com não ameríndios (em sua maioria de origem africana); e KAI apresenta média diferenciação e média taxa de mistura com populações não indígenas (tanto europeias quanto africanas) (TSUNETO *et al*, 2003; MARRERO *et al*, 2007). Estes achados reforçam a teoria migratória de indígenas com dispersão da Amazônia para a região sul e sudeste do Brasil (MARRERO *et al*, 2007).

Para GRC foi constatada diferenciação muito baixa com a população miscigenada brasileira de maneira geral, em concordância com o relatado por Marrero *et al* (2007) que encontrou cromossomos Y não ameríndios nas populações de Rio das Cobras (Guaranis M'Byá e Kaingangs), sugerindo misturas não nativas nestas populações. Estas misturas seriam provenientes principalmente de homens de ascendência europeia, o que justifica que a diferenciação de GRC, baixa com relação a todas as populações miscigenadas brasileiras, seja menor para com a população Sul e Sudeste (com maior imigração europeia) do que com as populações Norte e Nordeste (com maior imigração africana) (MARRERO *et al*, 2007).

#### 5.5.1.2 Structure

A análise de agrupamento realizada no programa STRUCTURE entre os cinco povos ameríndios e as cinco populações miscigenadas brasileiras, inferiu que o melhor número de agrupamentos (K) é K=2 (Delta K = 683.851516), um grupo representando os indígenas e outro representando a população brasileira. O agrupamento indica similaridade genética entre indígenas entre si e da população miscigenada brasileira entre si, evidenciando as diferenças entre estas (Figura 10). O compartilhamento mais evidente entre os dois grupos pode ser constatado na população norte, enquanto que as demais regiões brasileiras demonstram um compartilhamento menor e equivalente entre si. Os valores de Delta K, obtidos para diversas possibilidades de agrupamentos, se encontram disponibilizados no Anexo VI (Tabela 27).



**FIGURA 10** - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA ESTIMADO PARA OS INDIVÍDUOS E PARA AS POPULAÇÕES DE AMERÍNDIOS GUARANIS (GKW, GND E GRC), KAINGANGS (KRC E KIV) E POPULAÇÕES MISCIGENADAS BRASILEIRAS

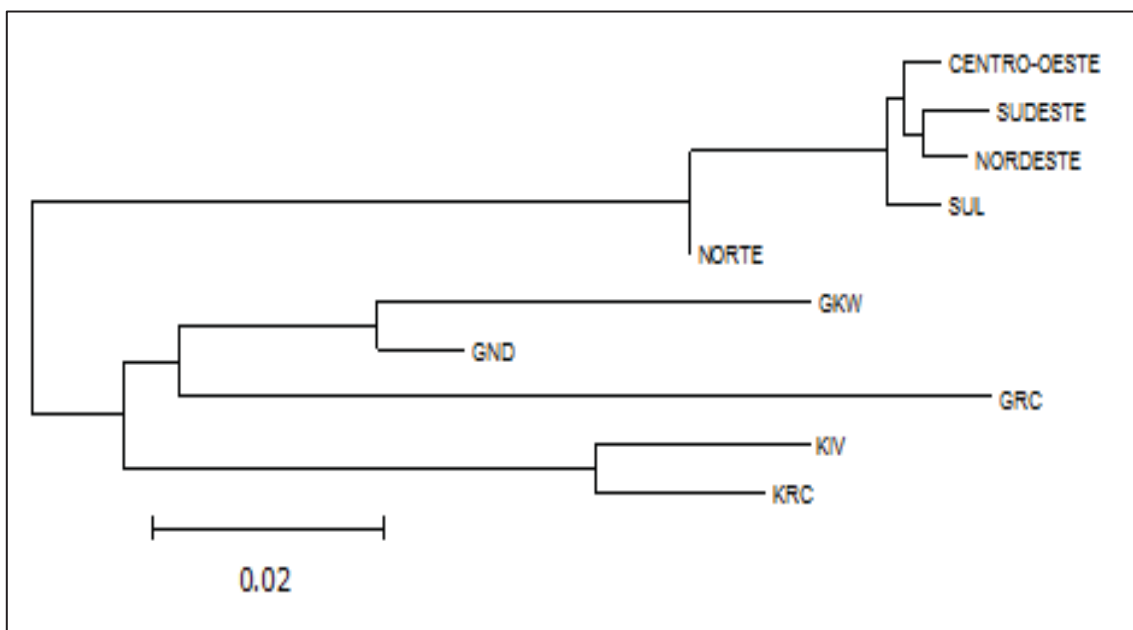
Considerando  $K=2$ . Cada indivíduo é representado por uma linha vertical com duas cores (verde e vermelho), em que a proporção de cada cor está relacionada às proporções de compartilhamento genético de cada um dos grupos populacionais inferidos.

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang do Ivaí; KRC: Kaingang de Rio das Cobras.

Fonte: O autor (2019).

### 5.5.1.3 Agrupamento hierárquico

O dendrograma demonstra que dois ramos distintos dão origem, respectivamente, aos indígenas e à população brasileira miscigenada (Figura 11). A população do Norte se encontra mais próxima das populações indígenas. Observa-se um distanciamento dos indígenas, como um todo, com relação às demais regiões brasileiras.



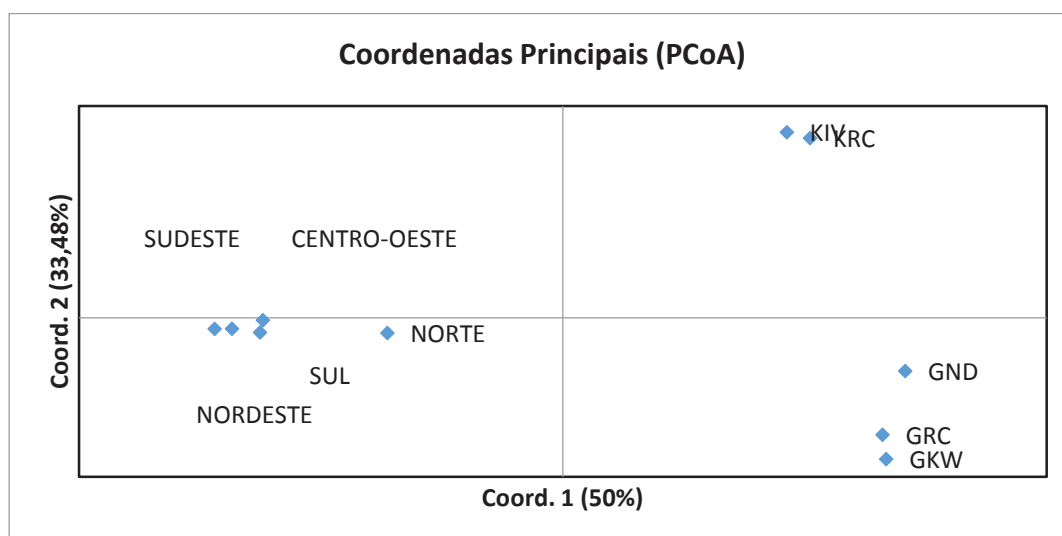
**FIGURA 11** - DENDROGRAMA GERADO COM BASE NOS MARCADORES 21 STRs, PARA AS POPULAÇÕES INDÍGENAS E BRASILEIRA MISCIGENADA.

Foi utilizado método de agrupamento neighbor-Joining e *bootstrap* de 1000 permutações. GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang do Ivaí; KRC: Kaingang de Rio das Cobras.

FONTE: O autor (2019).

#### 5.5.1.4 PCoA

O PCoA entre as populações ameríndias e miscigenada brasileira, utilizando as frequências alélicas dos 21 marcadores, demonstrou que a contribuição à variação dos componentes principais foi superior a 90%, quando considerados o primeiro, o segundo e o terceiro componentes. Os primeiros dois componentes contribuíram com 50% e 33,48%, respectivamente, e o terceiro, com 14,20%. Os indígenas são representados como grupos separados das populações brasileiras. Guaranis se encontram em um quadrante diferente do quadrante de Kaingangs (Figura 12).



**FIGURA 12** - PCOA DAS FREQUÊNCIAS ALÉLICAS DE 21 STRS TESTADOS PARA OS POVOS AMERÍNDIOS GUARANI E KAINGANG E AS POPULAÇÕES DE CINCO REGIÕES BRASILEIRAS.

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang do Ivaí; KRC: Kaingang de Rio das Cobras.

Fonte: O autor (2019).

Considerando-se que a ascendência ameríndia é, proporcionalmente, mais representada na região norte do Brasil, e de acordo com modelos propostos do povoamento indígena da América do Sul através de rotas costeiras seguidas de rotas pela porção norte do país (WANG *et al*, 2007), seria de se esperar, *a priori*, que a diferenciação entre os indígenas e esta região fosse menor, como foi constatado no caso de GKW e GND, pelo parâmetro Rst. A predominância da povoamento europeu na porção sul do país também justificaria a maior diferenciação verificada desta para com GKW e GND. Este resultado foi confirmado pelas análises de agrupamento do Structure e pelas análises realizadas pelo PCoA e pelo dendrograma, corroborando com a hipótese de origem amazônica para os povos Tupi-guarani e com a teoria de sua etnogênese na Amazonia, de onde teriam partido para o sul e para o leste (NEVES *et al*, 2011).

A população GRC se mostrou como uma população distinta e apartada das demais Guarani, tanto pela estatística R quanto pelas análises de dendrograma e PCoA, tendo inclusive demonstrado diferenciação, com relação à população brasileira, exatamente inversa: maior com relação à região Norte e Nordeste e menor com a região Sul. Sua localização geográfica, no Paraná,

difere da dos outros Guaranis (ambos do Mato Grosso do Sul) sugerindo padrões diferentes de migração e fluxo gênico, como relatado acima.

Os Kaingangs unificados mostraram-se mais diferenciados com relação ao Centro-Oeste e menos diferenciados com relação ao Nordeste.

#### 5.5.2 Populações Guarani e Kaingang, parentais e outros ameríndios

As heterozigosidades observadas nas três populações Guarani e nas duas populações Kaingang assemelham-se aos valores de heterozigosidade média dos M'byá-Guarani da Argentina (0,68) e de ameríndios do Brasil (Suruí, com 0,68 e Kairitiana, com 0,67), do México (0,66), da Guiana Francesa (0,67), Peru (0,66) e Alaska (0,69). A heterozigosidade obtida para GRC foi a menor observada entre todas as populações ameríndias consultadas (AMIGO *et al*, 2009; SALA *et al*, 2010, VILLA-LOBOS *et al*, 2013; CALLERARI-JACQUES *et al*, 2011; MAZIERES *et al*, 2010).

Dentre dez populações indígenas norte americanas, a heterozigosidade média (0,72) mostrou-se superior à dos indígenas Guarani e Kaingang. Somente o valor para os ameríndios Seri (0,64) e Esquimo (0,68) foi equivalente aos obtidos para Guarani e Kaingang (KANTHASWAMY *et al*, 2016; MCCULLOH *et al*, 2016; BUDOWLE *et al*, 2001).

As heterozigosidades médias dos ameríndios brasileiros Terena (0,74) e Apalai (0,72) se mostraram superiores à dos Guarani e Kaingang. Da mesma forma, os valores de heterozigosidade observados para os indígenas de El Salvador (0,77), Guatemala (0,74) e Gran Chaco na Argentina (0,73) foram superiores aos valores obtidos para os ameríndios do presente estudo (LOVO-GOMEZ *et al*, 2007; CARDOSO *et al*, 2015; MANTA *et al*, 2012; CROSSETI *et al*, 2008).

A heterozigosidade média de Guaranis (GKW, GND e GRC) e Kaingangs mostrou-se inferior à relatada para populações urbanas brasileiras, que variaram de 0,78 (região Norte) a 0,80 (nas regiões Sudeste e Nordeste); europeias, que variaram de 0,76 (na Itália) a 0,801 (Portugal) e 0,817 (Espanha); e no Japão (0,79) (GUSMÃO *et al*, 2016; BUDOWLE *et al*, 2001; CABEZAS SILVA *et al*, 2016; TIE *et al*, 2006; RODRIGUES *et al*, 2007).



Heterozigosidade elevada revela maior diversidade genética, apresentando relação positiva com o tamanho populacional e passível de ser perdida através de deriva genética e ou endocruzamentos (COLE, 1996; HAMRICK *et al*, 1991). Os resultados obtidos traduzem menor diversidade genética dos Guarani e Kaingang com relação a populações urbanas em geral, possivelmente em decorrência de constituírem populações menores, de menor tamanho efetivo, isoladas, e por estarem principalmente sujeitas à deriva, como relatado acima. Também evidenciam que a heterozigosidade tende a ser menor em ameríndios do que em populações urbanas, como já publicado. Observou-se que pode haver variações de amplitude de heterozigosidade dentro das populações indígenas, presumivelmente em decorrência da dinâmica populacional inerente ao histórico demográfico de cada uma. Isso é particularmente evidente na história de ondas migratórias e povoamento do continente americano, fortemente influenciados por segregação de natureza cultural e decorrentes de distanciamento geográfico.

Os valores obtidos de heterozigosidade para as demais populações ameríndias brasileiras avaliadas sugeriram valores semelhantes para Kaingang e Guarani do sul do Brasil e Kairitiana e Surui, e distintos de Terena e Apalai. Tais achados corroboram a hipótese do declínio de diversidade e de número de alelos nas populações ameríndias do leste da América do Sul, quando comparadas as populações do oeste (WANG *et al*, 2007).

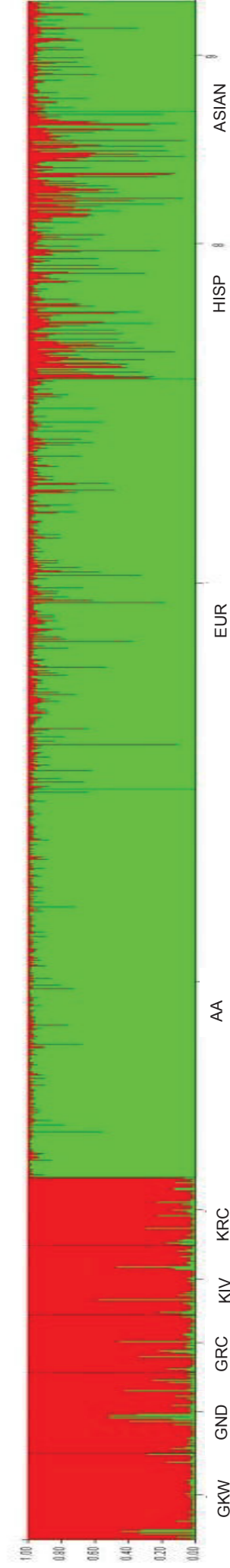
Os marcadores que apresentaram maior heterozigosidade diferiram conforme cada população ameríndia estudada (Tabela 8 e Anexo 2 – Tabela 21). Os marcadores com maior diversidade genética foram: SE33 (para GKW e KRC), D1S1656 (para GND e KIV) e FGA (para GRC). Na população de M'byá-Guarani da Argentina também constatou-se a maior heterozigosidade para FGA, juntamente com D16S539.

O marcador D2S441, com a menor heterozigosidade para as três populações Guarani (GKW, GND e GRC) é o que apresenta menor grau de diversidade gênica para estas populações. Já para KIV e KRC, observou-se que os marcadores com menor heterozigosidade foram D22S1045 e D3S1358, respectivamente.

### 5.5.2.1 Structure

A análise de agrupamento realizada no programa STRUCTURE entre Guaranis, Kaingangs e as quatro populações parentais miscigenadas, inferiu que o melhor número de agrupamentos (K) é K=2 (Delta K = 947.418510), um grupo representando os indígenas e outro representando a população parental miscigenada. O agrupamento evidencia a similaridade genética dos indígenas entre si, evidenciando sua diferença para com as populações parentais miscigenadas. Os valores de Delta K, obtidos para os números de agrupamentos, se encontram disponibilizados no Anexo VI (Tabela 28).

É visível um compartilhamento de similaridade genética dos dois grupos representados por Guaranis e Kaingangs e as populações parentais miscigenadas (Figura 13). GKW apresenta, dentre os indígenas, o compartilhamento menos prevalente. As populações parentais de origens predominantes hispânica e asiática demonstram o maior compartilhamento genético com indígenas Guarani e Kaingang.



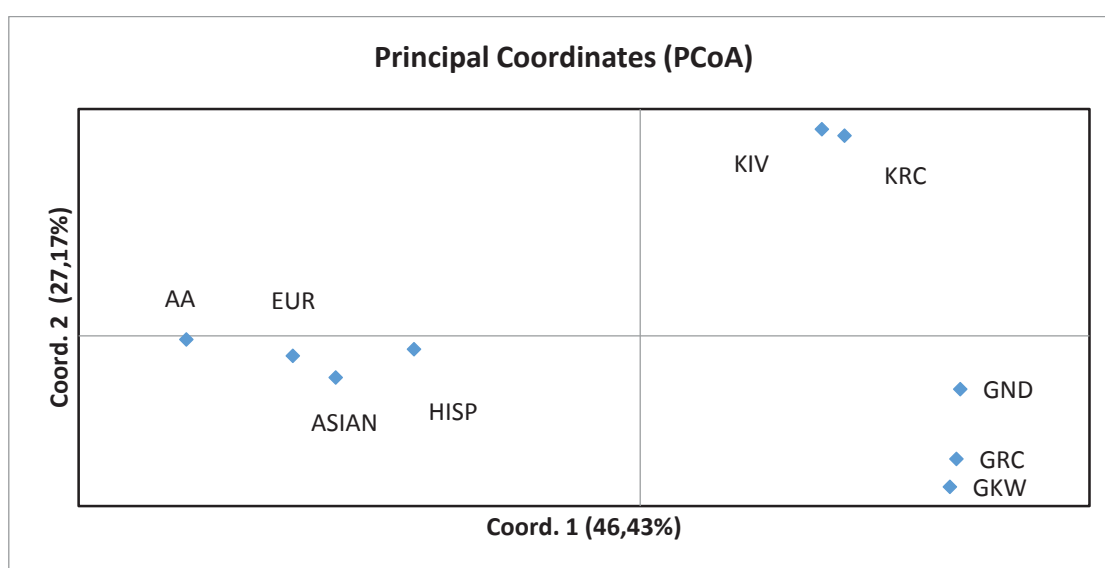
**FIGURA 13** - PROPORÇÕES DE COMPARTILHAMENTO DE SIMILARIDADE GENÉTICA PARA OS INDIVÍDUOS E POPULAÇÕES DE AMERÍNDIOS GUARANIS, KAINGANGS E POPULAÇÕES MISCIGENADAS PARENTAIS NORTE AMERICANAS COM PREDOMINÂNCIA AFROAMERICANA (AA), EUROPEIA (EUR), HISPÂNICA (HISP) E ASIÁTICA (ASIAN).

Considerando K=2. Cada indivíduo é representado por uma linha vertical com duas cores (verde e vermelho), em que a proporção de cada cor está relacionada às proporções de compartilhamento genético de cada um dos grupos populacionais inferidos.

FONTE: O autor (2019).

### 5.5.2.2 PCoA e PCA

O PCoA representando as relações entre as cinco populações indígenas e as populações parentais miscigenadas, utilizando as frequências alélicas dos vinte e um marcadores, demonstrou que a contribuição das coordenadas principais à variação foi de 86,30%, quando consideradas a primeira, a segunda e a terceira coordenada. As primeiras duas coordenadas contribuíram com 46,43% e 27,17%, respectivamente, e a terceira, com 12,71% (Figura 14).



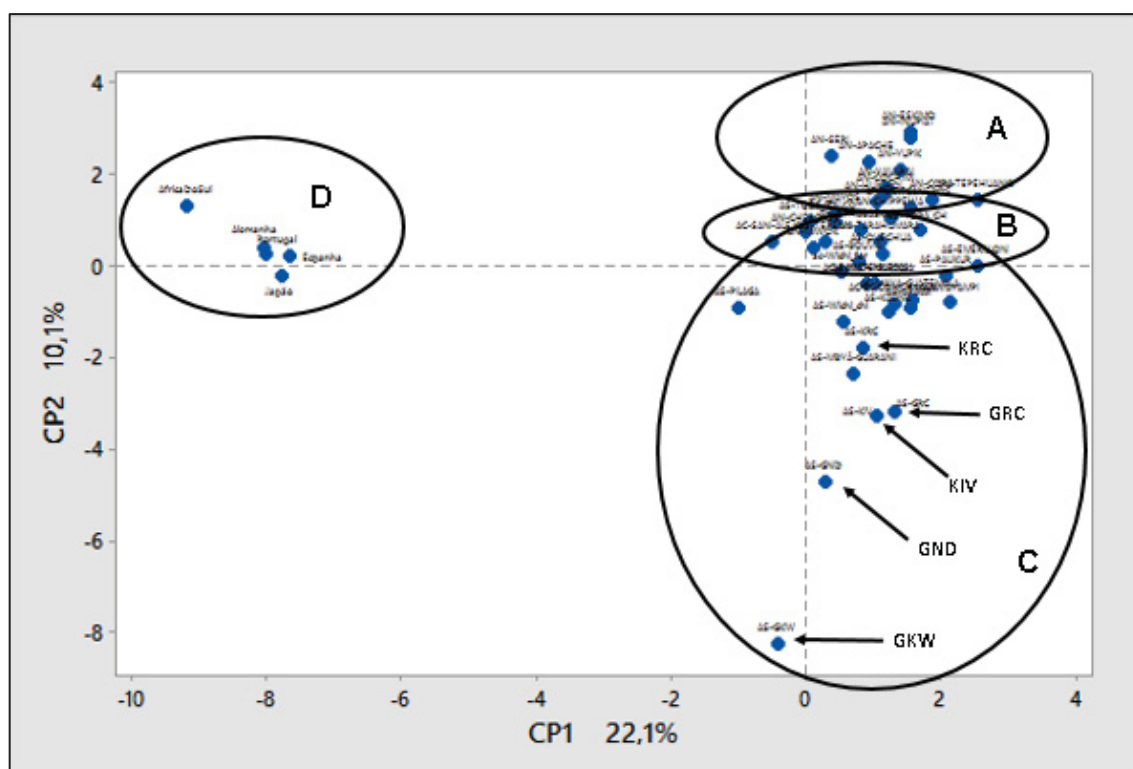
**FIGURA 14** - PCOA REALIZADA ENTRE TODAS AS 5 POPULAÇÕES AMERÍNDIAS E POPULAÇÕES PARENTAIS, PARA OS 21 MARCADORES STR.

AA: população parental miscigenada com predominância africana; CAUC: população parental miscigenada com predominância europeia; HISP: população parental miscigenada com predominância hispânica; ASIAN: população parental miscigenada com predominância asiática.

FONTE: O autor (2019).

O PCoA demonstra claramente três grupos distintos de acordo com suas características históricas e demográficas: populações parentais miscigenadas, Kaingangs e Guaranis. As populações parentais miscigenadas estão sistematicamente distantes das populações indígenas, demonstrando que são geneticamente muito diferentes. As populações Kaingang novamente são representadas em um grupo único e coeso, no mesmo quadrante, indicando grande similaridade genética. Os Guaranis são representados no mesmo quadrante (para as coordenadas principais 1 vs. 2), encontrando-se

GKW e GRC mais próximos entre si do que com relação a GND. Essa disposição entre os Guaranis muda conforme as coordenadas utilizadas, podendo aproximar GKW e GND, afastando GRC, ou separando estas três populações umas das outras. Entretanto, GKW e GND tendem a permanecer no mesmo quadrante em todas as representações (Figura 15).



**FIGURA 15** - PCA REALIZADO ENTRE TODOS OS AMERÍNDIOS E POPULAÇÕES PARENTAIS, PARA 6 MARCADORES STR (D3S1358, D8S1179, D21S11, D18S51, D19S539 e TH01).

Os valores em porcentagem indicam a parcela de variância explicada pelos respectivos componentes principais. CP1: Primeiro Componente Principal. CP2: Segundo Componente Principal. AS: América do Sul; AC: América Central; AN: América do Norte.

(A) População ameríndia da América do Norte; (B) População ameríndia da América Central; (C) População ameríndia da América do Sul; (D) Populações parentais.

Fonte: O autor (2019).

Através do PCA verificou-se que os componentes principais explicam 80% da variância encontrada entre estas populações, sendo o primeiro componente responsável por 22% da variação total, o segundo por 10% e o terceiro por 8%. O PCA gerado com os componentes principais 1 e 2 evidenciou que as cinco populações ameríndias são bastante diferenciadas com relação às populações do oeste, do centro e do norte das Américas, enquanto estas se apresentam sobrepostas e indiferenciadas (Figura 15).

Esta sobreposição das populações ameríndias das três Américas sugere similaridade genética entre os ameríndios. As populações ameríndias, entre si, apresentam realmente valores baixos a médios de  $F_{st}$  par a par, indicando diferenciação pequena e média. A região periférica do aglomerado, menos

concentrada, permite identificar populações ameríndias da América do Norte e da América do Sul distribuídas em polos opostos, de forma similar à sua localização geográfica, indicando menor similaridade genética.

Esta região periférica é melhor diferenciada pelo segundo componente (CP2), o qual distancia as cinco populações ameríndias em estudo do restante dos povos nativos americanos. A primeira população segregada é KRC, disposta próximo de M'byá-Guaranis da Argentina ( $F_{st}=0,048$ ), a partir da qual são representadas GRC ( $F_{st}=0,008$ ) e KIV ( $F_{st}=0,052$ ), afastando-se gradativamente mais das populações ameríndias sobrepostas. Kaingangs mostram-se menos diferenciados e mais similares com relação às populações indígenas norte e sul americanas. As populações GND e GKW são apresentadas de forma ainda mais separada das demais populações nativas. GND apresenta alta diferenciação ( $F_{st}=0,15$ ) com relação aos indígenas do Alaska, baixa a média diferenciação com relação a maioria dos nativos norte e sul americanos, média com nativos brasileiros Kairitiana e Surui ( $F_{st}=0,062$ ) e baixa com outros indígenas do Brasil ( $F_{st}=0,03$  com Apalai e  $F_{st}=0,04$  com Terena) e M'byá-Guaranis ( $F_{st}=0,02$ ). O CP2 separou GKW de todas as demais populações ameríndias.

Os resultados são similares aos obtidos por Wang *et al* (2007), com Guarani e Kaingang pertencendo a dois ramos diferentes e mais próximos a outros ameríndios da América do Sul e América Central, enquanto que mais distantes daqueles oriundos da América do Norte.

Esta análise é coerente com os modelos propostos de povoamento do leste da América do Sul e de redução da diversidade genética do norte para o sul das Américas e do oeste para o leste da América do Sul (WANG *et al*, 2007; TARAZONA *et al*, 2001). O PCA também sugere certa ausência de miscigenação dos povos parentais com os cinco povos ameríndios. Foi constatado, contudo, menor distanciamento entre os ameríndios as populações parentais originárias do Japão e da Espanha.

#### 5.5.2.3 Agrupamento hierárquico

O dendrograma obtido para estas populações, demonstra duas ramificações distintas e principais, uma originando as populações parentais e

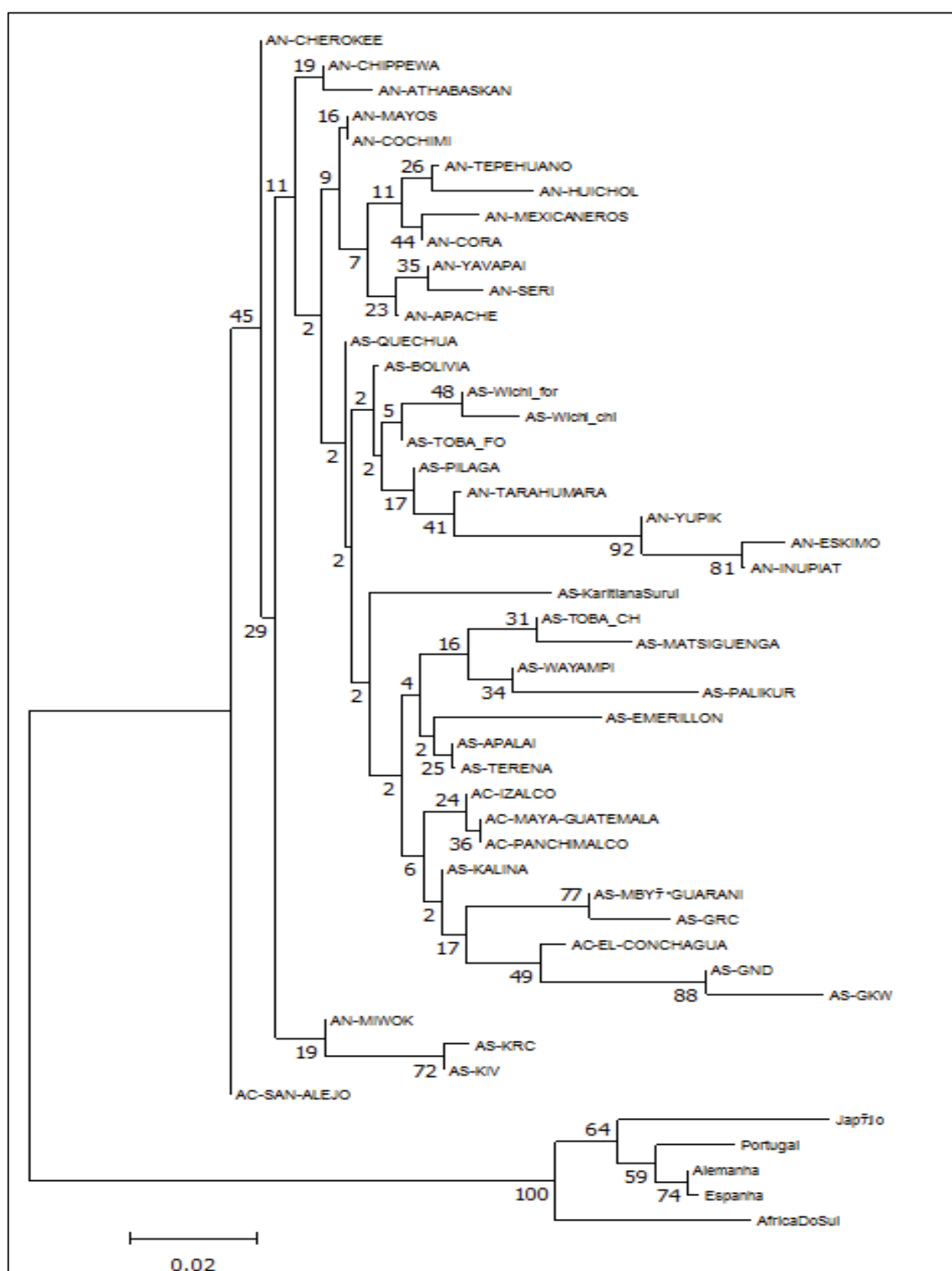
outra originando as populações ameríndias, indicando que se tratam de grupos geneticamente diferenciados, em correspondência ao evidenciado pela estatística F e por PCoA e PCA (Figura 16).

As ramificações que conduzem às diversas populações indígenas apresentam-se complexas e intrincadas. As duas populações Kaingang permanecem representadas próximas entre si e isoladas das demais populações indígenas em geral, com exceção de Miwok, população ameríndia da América do Norte. Diferentemente do PCA, KRC e M'Bya da Argentina não aparecem em ramos próximos, no dendrograma.

Os Guaranis aparecem representados a partir de um mesmo nó, que se divide em dois ramos. O primeiro ramo conduz à GRC e M'Byá-Guarani (Argentina), indicando similaridade genética entre estes. O outro conduz a dois outros ramos secundários, um levando à população Conchagua, de El Salvador e o outro levando à GKW e GND, corroborando com maior similaridade genética entre estas duas populações Guaranis e, menor, destas com relação à GRC. Estes resultados se encontram em conformidade aos obtidos por estatística F, análise de agrupamento, PCA e PCoA.

Dentre as populações parentais, a proximidade maior entre os cinco povos indígenas e as populações parentais miscigenadas com origem predominante asiática e com os denominados hispânicos, constatada no presente estudo, indica maior similaridade genética entre estas populações. As populações de origem predominante europeia e africana apresentam proporcionalmente menor similaridade genética com as cinco populações indígenas, como também constatado por Wang *et al* (2007) para outros indígenas da América do Sul. Estes achados estão de acordo com o que se sabe a respeito da colonização das américas: ameríndios descendem de pelo menos três correntes provenientes do leste asiático, sendo que o povoamento inicial seguiu uma expansão para o sul, pela costa, com pouco fluxo gênico posterior, especialmente na América do Sul (REICH *et al*, 2012).





**FIGURA 16** - DENDROGRAMA OBTIDO COM BASE EM 6 MARCADORES STR, PARA AS POPULAÇÕES AMERÍNDIAS E PARENTAIS.

AS: América do Sul; AN: América do Norte; AC: América Central. Foi utilizando o método de agrupamento Neighbor-Joining, para as populações ameríndias e parentais. Foi utilizado bootstrap de 10.000 randomizações. Em evidência: populações Guarani e Kaingang objeto do presente trabalho.

FONTE: O autor (2019).

#### 5.5.2.4 Estatística F

Entre Guaranis e Kaingangs e as populações parentais, foram constatados os valores mais elevados de diferenciação ( $0,169 < F_{st} < 0,219$ ) de todas as comparações populacionais realizadas, com valor médio de  $F_{st}=0,1937$ . O maior valor de  $F_{st}$  foi observado entre GRC e Japão. O menor, entre KIV e Alemanha (Anexo V – Tabela 25).

Considerando outras populações indígenas brasileiras e as populações Guaranis e Kaingangs, observou-se diferenciação média de Terena e Apalai ( $F_{st}=0,07$ ) com GKW e GRC e baixa diferenciação ( $F_{st}=0,036$  a  $0,040$ ) com GND, KIV e KRC. Todas as comparações mostraram diferenciação média ( $F_{st}=0,054$  a  $0,084$ ) para com Kairitiana e Surui (Anexo V – Tabela 25). Esta diferenciação condiz com o observado no dendrograma (Figura 16), no qual Kairitiana e Surui se encontram em ramificações mais distantes, com relação aos cinco povos ameríndios, do que Terena e Apalai.

Entre as três populações Guarani aqui estudadas e a população M'byá-Guarani da Argentina, seria, *a priori*, esperada baixa diferenciação, devido ao compartilhamento das origens culturais e linguísticas (SALA *et al*, 2010). A população M'byá-Guarani, da Argentina, apresentou realmente muito baixa diferenciação com relação a GRC ( $F_{st}=0,008$ ) e GND ( $F_{st}=0,022$ ); e média ( $R_{st}=0,052$ ), com relação a GKW.

Entre Kaingangs e Guarani-M'byá da Argentina se esperaria, também *a priori*, alta diferenciação, decorrente de suas diferenças linguísticas e culturais. A diferenciação constatada, entretanto, foi limítrofe entre baixa e média ( $F_{st}=0,05$ ). Este nível de diferenciação é intrigante, uma vez que a distância genética, estimada para quinze marcadores STR autossômicos, entre os Guaranis da Argentina e os Kaingangs, foi relatada como uma das mais elevadas quando comparada com a observada entre outras populações indígenas (SALA *et al*, 2010). Uma explicação possível poderia ser a utilização do parâmetro  $F_{st}$ , que por vezes pode subestimar o valor real de diferenciação. Este questionamento poderia ser melhor explorado através da determinação do valor do parâmetro de  $R_{st}$ . Para tanto, seria necessário que estivessem disponibilizadas na literatura as frequências genotípicas dos marcadores STR obtidas para a população Guarani-M'byá da Argentina.

Na América do Norte observam-se as únicas três ocorrências de alta diferenciação quando consideradas as populações ameríndias, todas com relação a Guaranis e a povos indígenas do Alasca, no extremo norte das Américas, sendo duas com relação à população Eskimo (GKW, com  $F_{st}=176$  e GND, com  $F_{st}=0,15$ ) e uma com relação à população Inupiat (GKW, com  $F_{st}=0,168$ ).

#### 5.5.2.5 Reconsiderações sobre os valores de $\theta$ para populações ameríndias

Um dos objetivos do cálculo do valor médio de  $R_{st}$  é a determinação do fator de correção  $\theta$  a ser utilizado nos cálculos forenses envolvendo populações específicas. Idealmente, cada população deveria ter seu valor de  $\theta$  estimado individualmente e foi neste sentido que as interpretações do valor de  $R_{st}$  foram feitas visando as cinco populações ameríndias brasileiras objeto deste estudo.

Todas as análises aqui conduzidas indicam a existência de diferenciação significativa entre os Guaranis, de forma que seria mais indicado aplicar o valor ( $R_{st}$ ) obtido individualmente para cada um deles: GKW (0,11), GND (0,06) e GRC (0,09). Entre Kaingangs, não tendo sido evidenciada diferenciação significativa, entende-se que ambas as populações podem ser consideradas, nos estudos forenses, como uma única população, sendo adotado o valor de  $R_{st}$  calculado para Kaingangs (0,08). O valor médio de  $R_{st}$  calculado para todas estas cinco populações ameríndias brasileiras em conjunto é de 0,09, sendo este o valor do fator de correção específico para as populações estudadas neste trabalho (Tabela 14).

Já os valores médios de  $F_{st}$  estimados para as cinco populações estudadas foram de 0,07 (para GKW), 0,05 (para GND), 0,08 (para GRC) e 0,06 para (Kaingangs), sendo o valor médio geral de 0,06 (Tabela 15).

**TABELA 14** – Valores médios de  $R_{st}$  e  $F_{st}$  para as 4 populações indígenas.

<i>População</i>	<i>R<sub>st</sub></i>	<i>F<sub>st</sub></i>
GKW	0,11	0,07

GND	0,06	0,05
GRC	0,09	0,08
KAI	0,08	0,06
VALOR MÉDIO	0,09	0,06

GKW: Guarani Kaiowá; GND: Guarani Ñandeva; GRC: Guarani M'byá; KAI: Kaingangs unificados.

FONTE: O autor (2019).

Os valores de  $R_{st}$  se mostraram, sistematicamente, superiores aos valores de  $F_{st}$ , para as cinco populações ameríndias.  $F_{st}$  pode de fato subestimar a diferenciação em populações estruturadas, como no caso em que se considerem marcadores STR, em virtude do padrão de mutação que estes apresentam. Neste caso, entende-se que o parâmetro  $R_{st}$  forneça estimativas mais precisas (BALLOUX *et al*, 2002). Entretanto, no presente estudo será necessário considerar os valores de  $F_{st}$  uma vez que, nas comparações realizadas com outras populações ameríndias, somente foi possível estimar o valor de  $F_{st}$ , conforme já explicado anteriormente.

Wang *et al* (2007), ao analisar 751 marcadores autossômicos (dos quais apenas alguns pertenceriam ao painel STR/CODIS) em 24 populações ameríndias, relata ter obtido um valor médio de  $F_{st}$  (0,081), o qual se assemelha ao  $R_{st}$  médio (0,09) obtido no presente estudo. Quando consideradas apenas as populações ameríndias da porção leste da América do Sul, foi estimado o valor de  $F_{st}$  de 0,147, superior ao valor médio encontrado no presente trabalho, o que pode ser devido ao maior número de populações geograficamente heterogêneas (WANG *et al*, 2007; YANG *et al*, 2010). É importante referir, entretanto, que tais valores foram estimados com base em marcadores STR, em sua grande maioria, não pertencentes ao painel do CODIS, tornando estes dados menos apropriados ao cálculo de  $\theta$  com propósito forense e a comparações com os dados aqui obtidos.

O valor médio de  $F_{st}$  estimado (0,06) para as cinco populações ameríndias brasileiras se mostrou superior aos recomendados pelo NRC (0,03) e aos estimados por Kanthaswamy *et al* (2018) e McCulloh *et al* (2016) (0,04) e por BUDOWLE *et al* (2001) (0,0282) para nativos norte-americanos, quando considerados treze marcadores STR/CODIS.

Apesar de sua provável origem comum, diferenças genéticas têm sido descritas entre os ameríndios, tendo sido constatado que suas populações apresentam os maiores níveis de diferenciação populacional quando comparadas a todas as outras populações distribuídas em territórios de tamanhos similares (KANTHASWAMY *et al*, 2016; KOHLRAUSH *et al*, 2005; WANG *et al*, 2007). Em um cenário ideal, portanto, a maioria das populações ameríndias deveria ser adequadamente representada em estudos abrangendo estimativas de diferenciação genética. Entretanto, admite-se que, na ausência de informações sobre diferenciação populacional a respeito de uma determinada população, dados de populações similares sejam utilizados (KANTHASWAMY *et al*, 2018; NRC II, 2016). A importância destes critérios é particularmente significativa ao se considerar estudos forenses que envolvam tais populações, nos quais se almeja utilizar o fator de correção ( $\theta$ ) mais representativo (KANTHASWAMY *et al*, 2016). Contudo, os bancos de frequências alélicas atualmente disponíveis apresentam representação limitada de ameríndios.

Investigações forenses normalmente carecem de informações precisas da origem populacional do perfil genético obtido para uma amostra questionada. Isso significa que não é comum que se disponha da informação de que um vestígio provenha especificamente de uma população indígena em particular. Amostras forenses costumam ser consideradas, portanto, oriundas da população local sem que se possa definir alguma particularidade, a não ser que esta seja apontada pela autoridade solicitante. Esta realidade torna mais coerente que o fator de correção seja definido de forma a incorporar uma maior representatividade das populações isoladas que compõem uma população geral estruturada.

Assim, foi considerado pertinente que se considerasse, juntamente as populações Guarani e Kaingang objeto do presente trabalho, os dados já disponibilizados para outras populações ameríndias. Foram então também estimados os valores de diferenciação genética entre populações ameríndias. Esta estimativa foi baseada somente no parâmetro  $F_{st}$  em virtude da categoria de dados que se encontra disponibilizada (frequências alélicas e não frequências genotípicas) para a maioria destas populações.

Considerando a heterogeneidade demográfica presente no continente americano foram portanto estimados os valores médios de  $F_{st}$  para as 43 populações ameríndias em conjunto e, também, quando consideradas separadamente, populações da América do Norte ( $N=17$ ), da América Central ( $N=5$ ) e da América do Sul ( $N=21$ ). Salienta-se que, dentre as populações da América do Sul, quatro populações ameríndias brasileiras estão incluídas (Apalai, Terena, Kairitiana e Surui), além das populações Guaraní e Kaingang. Os resultados obtidos apresentam valores similares entre si ( $F_{st}$  de 0,04 tanto para América do Sul quanto para a América do Norte). Para a América Central foi obtido um valor menor ( $F_{st}$  de 0,005), possivelmente em virtude do menor número de populações passíveis de serem comparadas. A média geral de  $F_{st}$ , obtida para todas as populações indígenas das Américas em conjunto, foi de 0,03, sendo este o valor do fator de correção que abrange todas as populações ameríndias estudadas (Tabela 15). Este valor se equipara ao recomendado pelo NRC (0,03) para populações isoladas.

No presente estudo, foi possível demonstrar, portanto, que o valor sugerido pelo NRC (1996), para o fator de correção ( $\theta$ ) para os cálculos forenses que envolvam populações isoladas, permanece adequado quando consideradas populações indígenas provenientes do continente americano. Sugere-se, todavia, que populações ameríndias adicionais sejam estudadas para que se disponha de mais informações individualizadas que possam incrementar os cálculos que visam a obtenção do fator de correção ( $\theta$ ) mais preciso para estas populações.

**TABELA 15** - Fator de Correção ( $\theta$ ) obtido a partir dos valores médios de  $F_{st}$  para as populações ameríndias.

<i><b>População Ameríndia</b></i>	<i><b><math>F_{st}</math> Médio - Fator de Correção (<math>\theta</math>)</b></i>
América do Norte	0,04
América Central	0,005
América do Sul	0,04
MÉDIA	0,03

FONTE: O autor (2019).

## 6 CONCLUSÕES

- Foram estimadas as frequências alélicas de 21 marcadores STR/CODIS para as populações indígenas Guarani (M'byá, Nandeva e Kaiowá) e Kaingang (de Rio das Cobras e do Rio Ivaí), dos estados do Paraná e Mato Grosso do Sul.

- Foi verificada a presença de desequilíbrio de ligação entre alguns pares de marcadores. Entretanto, a associação constatada nas populações Guarani e Kaingang foi de baixa magnitude e limitou-se a pares localizados em cromossomos distintos. Dessa forma, estes marcadores foram considerados não ligados, podendo ser tratados como locos independentes nas populações indígenas estudadas, para fins forenses.

- Entre Guaranis e Kaingangs verificou-se a existência de média a alta diferenciação populacional, indicando que há baixo fluxo gênico entre estas;

- Foi constatada baixa diferenciação e alta similaridade genética entre as duas populações Kaingang, indicando que, para análises forenses que utilizem o painel de marcadores STRs aqui utilizados, as frequências alélicas provenientes destas duas populações podem ser unificadas;

- Entre as três populações Guaranis a diferenciação populacional foi de média a alta, indicando que estas não podem ser reunidas para uma única estimativa de frequências alélicas;

- Foi verificada diversidade genética de baixa a média magnitude, entre Guaranis e Kaingangs, com relação à população brasileira miscigenada, e verificou-se que há um maior componente de ancestralidade indígena na população norte do país, quando comparada com outras regiões brasileiras;

- Observou-se redução de diversidade genética das populações Guarani e Kaingang situadas na região sudeste da América do Sul, quando comparadas às apresentadas pelos ameríndios da região oeste, da América Central e da América do Norte, corroborando modelos de migração do continente americano que preconizam a redução de heterozigosidade de norte a sul e de oeste ao leste das Américas.

- Estimou-se os parâmetros forenses propostos considerando o painel de marcadores estudado, os quais se mostraram adequados para serem utilizados nas populações Guaranis e Kaingangs estudadas.
- O valor médio de  $F_{st}$ , calculado considerando-se todas as populações indígenas, foi considerado o mais representativo da diferenciação genética das populações ameríndias e confirma a adequação do valor sugerido (0,03) para os cálculos forenses que envolvam populações indígenas das américas. Entretanto, recomenda-se novos estudos envolvendo populações ameríndias adicionais, para que se incorpore as informações representativas de mais populações aos cálculos forenses.



## REFERÊNCIAS

- AGUIAR, V.R; *et al.* New CODIS core loci allele frequencies for 96,400 Brazilian individuals. **Forensic Science International Genetics**, [online], v.13, p.e6-e12, nov. 2014. Acesso em: 02 set 2017.
- AMIGO J, Phillips C, Salas T, Fernández Formoso L, Carracedo A, Lareu M. pop.STR - An online population frequency browser for established and new forensic STRs. **Forensic Sci. Int. Gene. Suppl.** 2009. Disponível em [http://spsmart.cesga.es/about.php?dataSet=strs\\_local](http://spsmart.cesga.es/about.php?dataSet=strs_local) (10.1016/j.fsigss.2009.08.178). Acesso em jan/2019.
- AUGUSTO, D.G., *et al.* A deep look at KIR-HLA in Amerindians: Comprehensive meta-analysis reveals limited diversity of *KIR* haplotypes. **Human Immunology**. (2015), <http://dx.doi.org/10.1016/j.humimm.2015.01.025>. Acesso em 25.01.2019.
- BALLOUX, F., *et al.* The estimation of population differentiation with microsatellite markers. **Molecular Ecology**, v.11, p.155-65, 2002.
- BELTRAME, M.H. CD80 and CD86 polymorphisms in populations of various ancestries: 5 new CD80 promoter alleles. **Human Immunology**, v.73, p.111-117, jan. 2012.
- BUCKLETON, J.; *et al.* The effect of linkage on the calculation of DNA match probabilities for siblings and half siblings. **Forensic Science International**, p. 193-199, jul. 2006.
- BUDOWLE, B., *et al.* Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. **Int J Legal Med**, v.108, n.4, p.173-6. 1996.
- BUDOWLE, B., *et al.* Population Data on the Thirteen CODIS Core Short Tandem Repeat Loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. **Journal Of Forensic Sciences**, vol. 44, n.6, p.1277-86, nov. 1999.
- BUDOWLE, B., *et al.* Population Variation at the CODIS core short tandem repeat loci in Europeans. **Legal Medicine**, vol. 3, p.29-33, dez. 2000.
- BUDOWLE, B., *et al.* CODIS STR Loci Data from 41 Sample Populations. **Journal Of Forensic Sciences**, vol.46, n.3, p.453-89, mai,2001.
- BUDOWLE, B., *et al.* Population studies on three Native Alaska population groups using STR loci. **Forensic Sci Int**, vol.129, n.1, p.51-7, set,2002.

BUDOWLE, B., *et al.* Population studies on three Native Alaska population groups using STR loci. **Forensic Sci Int**, v.129, n.1, p.51-7, set, 2009.

BURGUETE-ARQUETA, N., *et al.* Forensic-paternity effectiveness and genetics population analysis of six non-CODIS mini-STR loci (D1S1656, D2S441, D6S1043, D10S1248, D12S391, D22S1045) and SE33 in Mestizo and Amerindian populations from Mexico. **Annals of Human Biology**, v.43, n.6, nov, 2016.

BUTLER, J.M.; *et al.* Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis. **Forensic Sciences Review**, v.24, (1), p.15-26, jan. 2012.

CABEZAS SILVA, R., *et al.* Analysis of 17 STR data on 5362 southern Portuguese individuals-an update on reference database. **Forensic Sci Int Genet**, 2016 Mar; 21:e10-2. doi: 10.1016/j.fsigen.2015.11.007. Epub 2015 Nov 22.

CALLEGARI-JACQUES, S.M.; *et al.* Autosome STRs in native South America-Testing models of association with geography and language. **Am J Phys Anthropol**. Vol.145, n.3, p.371-81, jul. 2011.

CARDOSO, S., *et al.* Analysis of 16 autosomal STRs and 17 Y-STRs in an indigenous Maya population from Guatemala. **Int J Legal Med**. V.130, n.2, p.465-6, mar, 2016.

CHAKRABORTY, R., WEISS, K.M. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. **Proceedings of the National Academy of Sciences of the United States of America**, vol.85, n. 23, p. 9119-23, dez. 1988.

CHAKRABORTY, R. Sample size requirements for addressing the population genetic issues of forensic use of DNA. **Human biology**, vol.64, n. 2, p. 141-59, abr. 1992.

CHEVITARESE, J. Determinação da Estrutura Genética das Populações Humanas e Inferência dos Fatores Evolutivos Que Contribuíram para sua Formação. 98F. Dissertação (Mestrado), Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 2009.

COLE, T.C. Genetic variation in rare and common plants. **Annu.Rev.Evol.Syst**. v.34. p.213-37, 2003.

COLLINS, A; *et al.* Likelihood ratios for DNA identification. **Proceedings of the National Academy of Sciences of the United States of America**, v. 91, n.13, p. 6007-11, jun. 1994.

CROSSETI, S.G., *et al.* Autosomal STR genetic variability in the Gran Chaco native population: homogeneity or heterogeneity? **Am J Hum Biol.** V.20, p.704-11, nov-dez.2008.

CURRAN J.M., TRIGGS,C.M., BUCKLETON, J., WEIR, B.S. Interpreting DNA mixtures in structured populations. **Journal of Forensic Science**, v.44(5), p.987-995, 1999.

CURRAN,J.M. *et al.* The appropriate use of subpopulation corrections for differences in endogamous communities. **Forensic Science International**, v.168 (2–3), p.106-11, mai, 2007.

EHLER, E; *et al.* Forensic genetic analyses in isolated populations with examples of Central European VALachs and Roma. **Journal of Forensics and Legal Medicine**, v. 48, p. 46-52, mai. 2017.

ELHAIK, E. Empirical Distributions of FST from Large-Scale Human Polymorphism Data. **PLoS One**, <online>, vol. 7, n. 11, e49837, nov. 2012. Acesso em 23 set 2017.

EARL, D. A., VONHOLDT, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, v. 4, n. 2, p.359-361, jul, 2012.

Evanno *et al.*, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Molecular Ecology** , v.14 , p.2611 - 2620. Disponível em <[http://users.soe.ucsc.edu/~dearl/software/struct\\_harvest](http://users.soe.ucsc.edu/~dearl/software/struct_harvest)>

EXCOFFIER, L.; *et al.* Arlequin (version 3.0): An integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online, Auckland**, v. 1, p. 47-50, 2005. Acesso em 15 set 2017.

FAGUNDES, N. J. R.; BONATTO, S. L.; CALLEGARI-JACQUES, S. M.; SALZANO, F. M. Genetic, geographic, and linguistic variation among South American Indians: possible sex influence. **American Journal of Physical Anthropology**, v. 117, n. 1, p. 68-78, 2002.

FAN, H., *et al.* **A Brief Review of Short Tandem Repeat Mutation.** Genomics Proteomics Bioinformatics, **V.5, n.1, p.7-14, jun, 2007.**

FRANCEZ, P.A. da C., *et al.* Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed populationof the Brazilian Amazon. **Genetics and Molecular Biology**, v.34, n.1, p.35-39, 2011.

GAGGIOTTI,O.E., *et al.* Comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. **Molecular Ecology**, v.8, p.1513-1520, 1999.

GASTALDO, A.Z., *et al.* Allele frequencies of 20 autosomal STR in the population from Rio Grande do Sul Southern Brazil. **Forensic Science International**, v.19, p.15-17, mai. 2015

GILL, A. P.; *et al.* Genotyping and interpretation of STR-DNA: low-template, mixtures and database matches - twenty years os research and development. **Forensic Sciences International Genetics**, vol. 18, p. 100-17, set. 2015.

GRATTAPAGLIA, D., *et al.* Brazilian population database for the 13 STR loci of the AmpFISTR Profier Plus and Cofiler multiplex kits. **Forensic Science International**, local, v.118, n.1, p. 91-94, abr. 2001.

GROVER, A., *et al.* Development and use of molecular markers: past and present. **Critical Reviews in Biotechnology**, Early Online, p.1–13.

GUO, S.W.; *et al.* Performing the exact test of Hardy-Weinberg proportion for multiple alleles. **Biometrics, Washington**, v. 48, n. 2, p. 361-372, 1992.

GUSMÃO, L.; *et al.* Population data of the 21 autosomal STRs included in the GlobalFiler kits in population samples from five Brazilian regions. **Forensic Science International: Genetics**, v.26, e28-230, 2017.

GUPTA, P. K.; RUSTGI, S.; KULWAL, P. L. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant molecular Biology*, Amsterdam, n. 57, p. 461- 485, jan. 2005.

HAMRICK, J.L., *et al.* Levels of allozyme diversity in populations of uncommon neotropical tree species. **Journal of Tropical Ecology**, v.7, p.395-99, 1991.

HARDY, G. H. (1908) Mendelian proportions in a mixed population. **Science** v.28, p.49-50, 1908.

HARES, D.R. Expanding the CODIS core loci in the United States. **Forensic Science International: Genetics**, <online>, v.6, p.e52–e54, jan.2012.

HILL, C.R., *et al.* U.S. population data for 29 autosomal STR loci. **Forensic Science International: Genetics**, v.7, e82-e83, mai.2013.

HOLSINGER, K. E.; *et al.* Genetics in geographically structured populations: defining, estimating and Interpreting FST. **Nature Reviews Genetics. Basingstoke**, v.10, p.639-50. Set, 2009.

HUTZ, M.H., *et al.* **Low levels of STRP variability** are not universal in American Indians. *Hum. Biol.* V.74, n.6, p.791-806, dez, 2002.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Censo Demográfico 2010. Disponível em: <<http://www.censo2010.ibge.gov.br>>. Acesso em 14 ago 2017.

INSTITUTO SOCIOAMBIENTAL. **Povos indígenas do Brasil**. Disponível em: < [https://pib.socioambiental.org/pt/P%C3%A1gina\\_principal](https://pib.socioambiental.org/pt/P%C3%A1gina_principal) >. Acesso em: 16 ago. 2017.

JIAN, T., *et al.* Genetic Polymorphisms of 15 STR Loci in a Japanese Population. **J Forensic Sci**, v.51, n.1, p.188-9, jan, 2006.

JOBLING, M.A.; GILL, P. Encoded evidence: DNA in forensic analysis. **Nature Reviews Genetics**, **Basingstoke**, v. 5, p.739-51, out. 2004.

JOHNSON, R. A.; WICHERN, D. W. Applied Multivariate Statistical Analysis. New Jersey, USA: Englewood Cliffs, 642p. 1992.

KANTHASWAMY, S., *et al.* Genetic and ethnohistoric evidence suggest current Native American population datasets in the FBI's CODIS Database are not sufficiently representative. **Forensic Sci Int: Genet.**, p.e13-5, mai. 2016.

KANTHASWAMY, S., *et al.* Native American population data based on the Globalfiler(®) autosomal STR loci. **Forensic Sci Int: Genet.** v.24, p.e12-e13, set. 2016.

KANTHASWAMY, S., *et al.* Assessing the FBI's Native American STR database for random match probability calculations. **Legal Medicine**. v.30, p.52-55, nov. 2017.

KLING, D., *et al.* Using Object Oriented Bayesian Networks to Model Linkage, Linkage Disequilibrium and Mutations between STR Markers. **PLoS One**, <online>, v.7, n.9, e43873. set. 2012. Acesso em: 23 set 2017.

KOHLRAUSCH, F., *et al.* Geography Influences Microsatellite Polymorphism Diversity in Amerindians. **American Journal of Physical Anthropology**, v.128, p.463-70, abr, 2005.

JLIU, Q.L.; *et al.* Potential of 13 linked autosomal short tandem repeat loci in pairwise kinship analysis. **Electrophoresis. Weinheim**, v.00, P.1-7, out. 2016.

LAHIRI, D.K.; NURNBERG, J.J.I. A rapid non enzymatic method for the preparation of HMW DNA from blood for RFLP studies. **Nucleic Acids Research**, v.19, p.19, 1991.

LANCASTER, A.; NELSON, M.P.; SINGLE, R.M.; MEYER, D.; THOMSON, G. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. **Pacific Symposium on Biocomputing**, v.8, pp.514-525, 2003.

LANGE, K. Match Probabilities in Racially Admixed Populations. **American Journal of Human Genetics**, v.52,n.2, p.305-311, fev. 1993.

LECLERCQ, S., *et al.* DNA Slippage Occurs at Microsatellite Loci without Minimal Threshold Length in Humans: A Comparative Genomic Approach. **Genome Biology and Evolution**, v. 2, p. 325-35, jan, 2010.

LEITE, F.P.N., *et al.* Linkage Disequilibrium Patterns and Genetic Structure of Amerindian and Non-Amerindian Brazilian Populations REvealed by Long-Range X-STR Markers. **American Journal of Physical Anthropology**, v.139, p. 404-12, jul, 2009.

LEWONTIN, R. C. On measures of gametic disequilibrium. *Genetics*, Heidelberg, v. 120, n. 3, p. 849-852, 1988.

LEWONTIN, R. C. The interaction of selection and linkage. I. General considerations; Heterotic Models. *Genetics*, Heidelberg, n. 49, p. 49-67, jan. 1964

LOLO-GÓMEZ, J., *et al.* Microsatellite autosomal genotyping data in four indigenous populations from El Salvador. **Forensic Sci Int**, v.170, n.1, p.86-91, jun, 2006.

MANIATIS, T.; FRITSCH, E.F.; SAMBROOK, J. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory, 1982.

MANTA, *et al.* Terena Amerindian group autosomal STR data: comparison studies with other Brazilian populations. **Mol Biol Rep**, 39:4455-4459, 2012.

MARTINEZ-GONZALEZ, L.J., *et al.* Characterisation of genetic structure of the Mayan population in Guatemala by autosomal STR analysis. *Annals of Human Biology*, v.43:5, p. 457-68, ago, 2015.

MAZIERES, S., *et al.* French Guiana Amerindian demographic history as revealed by autosomal and Y-chromosome STRs. **Ann Hum Biol**. V.38, n.1, p.76-883, jan, 2011.

MARRERO, A.R.; *et al.* Demographic and evolutionary of the Guarani and Kaingang natives of Brazil. *American Journal of Physical Anthropology*, New Jersey, n.132, p.301–310, 2007. National Research Council II Report. The evaluation of forensic evidence, **National Academy Press**, Washington, DC (1996).

MCCULLOH, *et al.* The genetic structure of native Americans in North America based on the Globalfiler® STRs. **Leg Med**, v.23, p.49-54, nov. 2016.

Minitab 17 Statistical Software, 2010. Computer software: State College, PA. Disponível em: <www.minitab.com>.

National Research Council (NRC), *The Evaluation of Forensic DNA Evidence*, **The National Academies Press**, Washington, DC, 1996.



NEVES, W.A., *et al.* Origem e dispersão dos Tupiguarani: o que diz a morfologia craniana? **Bol. Mus. Para. Emílio Goeldi. Ciên. Hum.**, v.6, n.1, jan/abr, 2011.

NORDBORG, M., *et al.* Next-generation genetics in plants. **Nature**, v. 456, p.720-3, dez, 2008.

PETZL-ERLER, M.L.; *et al.* The HLA polymorphism of two distinctive South American indian tribes: The Kaingang and the Guarani. **Tissue Antigens**, v. 41, p. 227-237, 1993.

Phillips C., *et al.* Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel. **Forensic Sciences International Genetics**, v.5, p.155–169, 2010.

OLIVEIRA, A.E., *et al.* Population data of 16 autosomal STR loci of the Powerplex ESX 17 System in a Brazilian Population from the State of São Paulo. **Forensic Sciences International Genetics**, v.11, p.15-7. Jul. 2014.

OLIVEIRA, I.D.B. **Desequilíbrio de ligação e análise de seleção genômica em cana-de-açúcar**. 85f. Dissertação (Mestrado), Escola de Agronomia, Universidade Federal de Goiás, fev, 2014.

OLIVEIRA, L. A., *et al.* High frequencies of alleles mica\*020 and mica\*027 in amerindians and evidence of positive selection on exon 3. **Nature, Genes and Immunity**, v. 9, 697–705, 2008.

PEAKALL, R. *et al.* GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. **Molecular Ecology Notes**. V.6, p.288-295, 2006.

PARHAM, P.; *et al.* Episodic evolution and turnover of HLA-B in the indigenous human populations of the Americas. **Tissue Antigens**, v.50, p.219-232, mai. 1997.

PAYSEUR, B.A., *et al.* Linkage Disequilibrium between STRPs and SNPs across the Human Genome. **Am. J. Hum. Genet.**, v.82, n.5, p.1039-50, mai, 2008.

PEAKALL, R. *et al.* GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. **Bioinformatics**, 28, 2537-2539, 2012.

PETTENER, D., PASTOR, S. TARAZONA-SANTOS, E. Surnames and genetic structure of a high-altitude Quechua community from the Ichu River Valley, Peruvian Central Andes, 1825-1914. **Hum. Biol.**, v.70, n.5, p.865-87, out, 1998.

PETZL-Erler, M.L., *et al.* The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. *Tissue Antigens*, v. 41, p.227-237, 1993.

POIARES, L., *et al.* Allele frequencies of 15 STR in a representative sample of the Brazilian population. **Forensic Sciences International Genetics**, v.4, p.61-3, mai. 2009.

PRITCHARD J. K., STEPHENS M. & DONNELLY P. Inference of population structure using multilocus genotype data. **Genetics** v.155, p.945–959, 2000.

RANGEL-VILLALOBOSA, H., *et al.* Forensic parameters for 15 STRs in eight Amerindian populations from the north and west of Mexico. **Forensic Science International: Genetics**. v.7, n.3, p.e62-e65, mai, 2013.

RAYMOND, M., ROUSSET, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. **J. Heredity**, 86:248-249, 1995.

REICH, D., *et al.* Reconstructing Native American population history. *Nature*, v. 488, p. 370–374, ago, 2012.

RISTOW, P., *et al.* Forensic statistics analysis toolbox (FORSTST): A streamlined workflow for forensic statistics. **Forensic Science International: Genetics Supplement Series**, dec, 2017.

RYCKMAN, K., *et al.* Calculation and use of the Hardy-Weinberg model in association studies. **Current Protocols in Human Genetics**, c..1, unid 1.18, abr. 2008.

ROCABADO, O., *et al.* Population genetic data for 15 STR loci (Identifiler™ kit) in Bolivia. **Legal Medicine**, v. 11, n. 6, p.302-304, nov, 2009.

RODRIGUES, E.L., *et al.* Allele frequencies data and statistic parameters for 13 STR loci in a population of the Brazilian Amazon Region. **Forensic Sciences International**. v.168, p.244-7, jun.2007.

RODRIGUES, E.L., *et al.* Genetic data on 15 STR autosomal loci for a sample population of the Northern Region of the State of Rio de Janeiro, Brazil. **Forensic Sciences International Genetics**. v.4, n.1, p.25-6, dec.2009.

RODOVALHO, R.G., *et al.* Development of a polymorphic short tandem repeat locus multiplex system for efficient human identification. **Genetics and Molecular Research**, <online>, v.16, n.2, gmr16029557, abr. 2017. Acesso em: 23 set 2017.

ROUSSET, F., 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. **Mol. Ecol. Resources** 8: 103-106.



SAITOU, N., NEI, M. The *Neighbor-Joining* method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v.4, p.406-425, 1987.

SALA, A., *et al.* Genetic analysis of six communities of M'byá-Guaraní inhabiting northeastern Argentina by means of nuclear and mitochondrial polymorphic markers. **Hum Biol**, v.82, n.4, p.433-56, ago, 2010.

SCHNEIDER, P.M. Scientific standards for studies in forensic genetics. **Forensic Science International**, v.165, p.238-243, jul, 2007.

Scientific Working Group on DNA Analysis Methods (SWGDAM): Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. United States 2010.

SILVA, D.A., *et al.* Statistical analyses of 14 short tandem repeat loci in Brazilian populations from Rio de Janeiro and Mato Grosso do Sul states for forensic and identity testing purposes. **Forensic Science International**. v.139, n.2-3, p.173-6, jan. 2004.

SILVA, N.M.; *et al.* Human neutral genetic variation and forensic STR data. **PLoS One**, v. 7, n. 11, p.49666, nov. 2012.

SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v. 139, n.1, p.457-62, jan, 1995.

STRIDER - Projeto do Institute of Legal Medicine, Medical University of Innsbruck, Áustria: STRs for identity ENFSI Reference database v.2 - disponível em <https://strider.online/>. Acesso em jan/2019.

SZPIECH, A.Z. *et al.* On the size distribution of private microsatellite alleles. **Theor Popul Biol**, v.80, n.2, p.100-113, set, 2011.

Scientific Working Group on DNA Analysis Methods (SWGDAM): Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. United States 2010.

TAKEZAKI, N., NEI, M., TAMURA K. POPTREEW: Web version of POPTREE for constructing population trees from allele frequency data and computing other population statistics. *Molecular Biology and Evolution*, v31, p.1622-1624, 2014. Disponível em < <http://www.med.kagawa-u.ac.jp/~genomelb/takezaki/poptreew/> >.

TAMURA, T.; *et al.* Evaluation of advanced multiplex short tandem repeat systems in pairwise kinship analysis. **Legal Medicine** (Tokio). Amsterdam, v.17, n.5, p.320-5, set, 2015.

TENDEIRO, J.N. 2005. Comparação de Dendrogramas: Obtenção de Distribuições Empíricas de Alguns Coeficientes. 206f. Dissertação (Mestrado

em Estatística Aplicada e Modelação) – Faculdade de Ciências da Universidade do Porto, Faculdade de Engenharia da Universidade do Porto, Porto, 2005.

TEREBA, A. 1999. Tools for analysis of population statistics, **Profiles DNA**.n.3. p.14–16.

TIE, J., *et al.* Genetic Polymorphisms of 15 STR loci in a Japanese population. **J Forensic Sci.** v.51, n.1, p.188-9, jan, 2006.

TILLMAR, A.O., *et al.* Evaluation of the impact of genetic linkage in forensic identity and relationship testing for expanded DNA marker sets. **Forencis Science International: Genetics**, v.26, p.58-65, out, 2015.

TRACEY, M. Short tandem repeat-based identification of individuals and parents. **Croatian Medical Journal**, <online>, v.42, p.233-238, 2001. Disponível em <<http://www.cmj.hr/2001/42/3/11387629.pdf>>. Acesso em:10 se /2017.

TSUNETO, L.T., *et al.* HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. **Tissue antigens**, v.62, n.6, p.512-26, dez. 2003.

TVEDEBRINK, T. Overdispersion in allelic counts and -correction in forensic genetics.

Theoretical Population Biology, v.78, p.200-210, 2010.

VULLO, C., *et al.* Genetic polymorphism of 22 autosomal STR markers in Paraguay. **Forensic Science International: Genetics**, v.25, e16-e17, ago. 2016.

WADT, L.H.de O., *et al.* Estrutura genética e sistema de acasalamento de *Piper hispidinervum*. **Pesq. Agropec. Bras**, v.39, n.2, p.151-7, fev. 2004.

WALSH, P.S., *et al.* Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. **Nucleic Acids Research**, vol. 24, no. 14, p.2807–2812, mai, 1996.

WANG, S., *et al.* Genetic variation and population structure in native Americans. **PLoS Genetics**, v.3, n.11, p.185, nov, 2007.

WANG, R., *et al.* When genetic distance matters: Measuring genetic differentiation at microsatellite loci in whole-genome scans of recent and incipient mosquito species. **PNAS**, v.98, n.19, p.10769-774, St, 2001.

WAPLES, R. S.; ENGLAND, P. R. Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. **Genetics**, Heidelberg, v. 189, n. 2, p. 633-644, 2011.

WEIR, B.S., COCKERHAM, C.C. Estimating F-Statistics for the Analysis of Population Structure. **Evolution**, v.38, p.1358-70, nov, 1984.

WEISSTEIN, E.W. Bonferroni Correction. MathWorld - A Wolfram Web Resource. Disponível em: <<http://mathworld.wolfram.com/BonferroniCorrection.html>> Acesso em: 25 jan. 2019

WITTKÉ-THOMPSON J. K., PLUZHNIKOV A., COX N. J. (2005) Rational inferences about departures from Hardy-Weinberg equilibrium. The American Journal of Human. **Genetics** 76:967-986. ANO

Wright, S. The interpretation of population structure by F-statistics with special regard to systems of mating. **Evolution**, Madison, v.19, p.395-420, mai.1965.

WEINBERG W. (1908) Über den Nachweis der Vererbung beim Menschen. **Jh. Ver. vaterl., Naturk. Württemb.** v.64, p.369–382. 1908.

YANG, N.N., *et al.* Contrasting Patterns of Nuclear and mtDNA Diversity in Native American Populations. **ANNALS OF HUMAN GENETICS**, V.74, P.525-38, set, 2010.

ZAPATA, C. The D'measure of overall gametic disequilibrium between pairs of multiallelic loci. **Evolution, Massachusetts**, v. 54, n. 5, p. 1809-1812, 2000.

ZARE ZARE CHAHOUKI, M.A. Multivariate analysis techniques in environmental science. **Earth and Environmental Sciences**, Imran Ahmad Dar and Mithas Ahmad Dar, IntechOpen, dez, 2011. Disponível em <<https://www.intechopen.com/books/earth-and-environmental-sciences/multivariate-analysis-techniques-in-environmental-science>>

ZHIVOTOVSKY, L.A. *et al.* Developing STR databases on structured populations: The native South Siberian population versus the Russian population. **Forensic Science International: Genetics**, v.3, n.4, e111-116, 2009.

## ANEXO I

**TABELA 16-** Frequências alélicas da população Guarani Kaiowá (GKW).

[illegible]

Pop	Allele/n	D3S1358	TH01	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	SE33	VWA	D8S1179	TPOX	FGA	D2S1338	D19S433	D10S1248	D22S1045	D151656	D2S441	D12S398
GND	5	71	71	71	71	71	71	71	71	70	71	71	71	71	71	71	71	71	71	71	71	71
	N													0,000								
	6		0,606															0,000				
	7		0,303																			
	8		0,014																			
	9		0,014																			
	10				0,000	0,106	0,077	0,007	0,014	0,000			0,014	0,282						0,028	0,739	
	11				0,000	0,035	0,021	0,085	0,049	0,143			0,042	0,021				0,000	0,007	0,000	0,162	
	12				0,000	0,444	0,190	0,486	0,183	0,286			0,000	0,486			0,000	0,021	0,021	0,028	0,028	
	13	0,000			0,218	0,042	0,225	0,331	0,197	0,500	0,014		0,070	0,190			0,204	0,120	0,077	0,007	0,007	
	14	0,077			0,190	0,035	0,092	0,056	0,085	0,071	0,014		0,239	0,190			0,183	0,430	0,120	0,056	0,056	
	15	0,493			0,366		0,077		0,000			0,007	0,401				0,204	0,338	0,641	0,155		
	16	0,239			0,077						0,035	0,000	0,218			0,000	0,085	0,092	0,310	0,035	0,000	
	17	0,134			0,120				0,021		0,190	0,380	0,014			0,324	0,085	0,092	0,042	0,035	0,014	
	18	0,049			0,007				0,000		0,120	0,437			0,007	0,077	0,141	0,099	0,641	0,155	0,070	
	19	0,007			0,000						0,141	0,042	0,134		0,007	0,141	0,127	0,338	0,190	0,028	0,423	
	20				0,000						0,077	0,000			0,000	0,120	0,127	0,155	0,310	0,035	0,373	
	21				0,000					0,028					0,070	0,014	0,127	0,155	0,310	0,035	0,063	
	22				0,000										0,042	0,190	0,127	0,155	0,310	0,035	0,063	
	23				0,000										0,211	0,007	0,127	0,155	0,310	0,035	0,028	
	24														0,204	0,014	0,127	0,155	0,310	0,035	0,021	
	25														0,190	0,014	0,127	0,155	0,310	0,035	0,007	
	26														0,007		0,127	0,155	0,310	0,035	0,007	
	27			0,000											0,007		0,127	0,155	0,310	0,035	0,007	
	28			0,000							0,000				0,007		0,127	0,155	0,310	0,035	0,007	
	29			0,120											0,141		0,127	0,155	0,310	0,035	0,007	
	30			0,472											0,000		0,127	0,155	0,310	0,035	0,007	
	31			0,014											0,000		0,127	0,155	0,310	0,035	0,007	
	36			0,000											0,007		0,127	0,155	0,310	0,035	0,007	
	53	0,000													0,007		0,127	0,155	0,310	0,035	0,007	
	63	0,000													0,007		0,127	0,155	0,310	0,035	0,007	
	93	0,063													0,007		0,127	0,155	0,310	0,035	0,007	
	113														0,007		0,127	0,155	0,310	0,035	0,007	
	122														0,007		0,127	0,155	0,310	0,035	0,007	
	132														0,007		0,127	0,155	0,310	0,035	0,007	
	142														0,007		0,127	0,155	0,310	0,035	0,007	
	152														0,007		0,127	0,155	0,310	0,035	0,007	
	153														0,007		0,127	0,155	0,310	0,035	0,007	
	162														0,007		0,127	0,155	0,310	0,035	0,007	
	163														0,007		0,127	0,155	0,310	0,035	0,007	
	173														0,007		0,127	0,155	0,310	0,035	0,007	
	183														0,007		0,127	0,155	0,310	0,035	0,007	
	203														0,007		0,127	0,155	0,310	0,035	0,007	
	242														0,007		0,127	0,155	0,310	0,035	0,007	
	252														0,007		0,127	0,155	0,310	0,035	0,007	
	262														0,007		0,127	0,155	0,310	0,035	0,007	
	282														0,007		0,127	0,155	0,310	0,035	0,007	
	292														0,007		0,127	0,155	0,310	0,035	0,007	
	302			0,007											0,007		0,127	0,155	0,310	0,035	0,007	
	312			0,077											0,007		0,127	0,155	0,310	0,035	0,007	
	322			0,176											0,007		0,127	0,155	0,310	0,035	0,007	
	332			0,134											0,007		0,127	0,155	0,310	0,035	0,007	
	432														0,007		0,127	0,155	0,310	0,035	0,007	

ABELA 47 - Frequências alelicas da população Guarani Nandeva (GND).

Pop	Allele/n	D3S1358	TH01	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	SE33	VWA	D8S1179	TPOX	FGA	D2S1338	D19S433	D10S1248	D22S1045	D1S1656	D2S441	D12S391
GRC	N	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	50
5																						
6			0.343																			
7			0.294																			
8			0.000																			
9			0.000																			
10																						
11																						
12		0.000																				
13																						
14		0.000																				
15		0.333																				
16		0.559																				
17		0.039																				
18		0.049																				
19		0.020																				
20																						
21																						
22																						
23																						
24																						
25																						
26																						
27																						
28			0.000																			
29			0.010																			
30			0.265																			
31			0.363																			
32			0.137																			
36			0.000																			
53			0.000																			
63			0.000																			
93			0.363																			
113																						
122																						
132																						
142																						
152																						
153																						
162																						
163																						
173																						
183																						
203																						
242																						
252																						
262																						
272																						
282																						
292																						
302																						
312			0.000																			
322			0.029																			
332			0.167																			
332			0.029																			
351																						
432																						

TABELA 18 - Frequências alélicas da população Guarani M'byá (GRC).







Pop	Alleles	D3S1358	TH01	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	SE33	vWA	D8S1179	TPOX	FGA	D2S1338	D19S433	D10S1248	D22S1045	D151656	D2S441	D12S391
KRC	N	60	60	60	59	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
	5														0,008							
	6		0,108																			
	7		0,458																			
	8		0,000																			
	9		0,017																			
	10				0,000	0,058	0,000	0,025	0,000	0,367	0,000		0,000		0,408					0,000	0,383	
	11				0,034	0,508	0,075	0,308	0,117	0,242			0,058		0,050				0,025	0,000	0,533	
	12	0,000			0,042	0,125	0,092	0,483	0,325	0,200			0,033		0,242			0,008	0,008	0,000	0,025	
	13				0,203	0,025	0,208	0,075	0,183	0,533	0,000		0,067		0,275			0,125	0,083	0,158	0,000	
	14	0,000			0,424		0,050	0,075	0,008	0,025		0,025	0,308					0,033	0,550	0,158	0,058	
	15	0,667			0,017				0,000		0,000	0,000	0,150					0,333	0,308	0,100		
	16	0,275			0,153						0,000	0,567	0,000					0,000	0,017	0,100		
	17	0,058			0,034						0,125	0,392						0,225	0,392	0,083		
	18	0,000			0,076						0,058	0,017						0,050	0,008	0,033		
	19	0,000			0,000						0,050	0,000						0,083	0,017	0,000		
	20										0,108	0,000						0,017	0,083			
	21										0,008							0,167	0,017	0,233		
	22										0,008							0,133	0,092	0,033		
	23										0,000							0,050	0,242	0,092		
	24																	0,283	0,000	0,000		
	25																	0,267	0,000			
	26																	0,042				
	27			0,033							0,000								0,008			
	28		0,083																			
	29		0,058																			
	30		0,483																			
	31		0,025																			
	36		0,000																			
	53		0,000																			
	63		0,000																			
	93		0,417																			
	113																					
	122																					
	132																					
	142																					
	152																					
	153																					
	162																					
	163																					
	173																					
	183																					
	203																					
	242																					
	252																					
	262																					
	272																					
	282																					
	292																					
	302			0,000																		
	312			0,025																		
	322			0,092																		
	332			0,175																		
	351			0,025																		
	432																					

**TABELA 20** - Frequências alélicas da população Kaingang – Rio das Cobras (KRC).

## ANEXO II

TABELA 21 - Parâmetros forenses estimados por população ameríndia.

Pop	Alelos	D3S1358	TH01	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	SE33	vWA	D8S1179	TPOX	FGA	D2S1338	D19S433	D10S1248	D22S1045	D151656	D25441	D12S591
GKW	N	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76
	2N	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152
	Na	5	7	7	8	5	7	6	6	5	16	5	5	4	8	9	9	4	5	11	5	9
	Ne	3,131	2,203	3,776	4,604	2,098	4,026	2,796	2,233	2,804	9,675	3,413	3,146	2,768	5,027	5,192	4,111	2,671	4,186	1,552	3,741	3,741
	FAM	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033	0,033
	Ho	0,763	0,421	0,737	0,763	0,539	0,697	0,711	0,513	0,684	0,882	0,684	0,776	0,605	0,816	0,816	0,803	0,566	0,605	0,711	0,382	0,763
	He	0,681	0,546	0,735	0,783	0,523	0,752	0,642	0,552	0,643	0,897	0,707	0,682	0,639	0,801	0,807	0,757	0,626	0,547	0,761	0,356	0,733
	PC	0,179	0,276	0,116	0,088	0,300	0,098	0,217	0,246	0,204	0,028	0,139	0,155	0,192	0,076	0,078	0,094	0,204	0,348	0,086	0,429	0,122
	l emt	5,6	3,6	8,6	11,3	3,3	10,2	4,6	4,1	4,9	35,2	7,2	6,4	5,2	13,1	12,9	10,7	4,9	2,9	11,6	2,3	8,2
	PD	0,821	0,724	0,884	0,912	0,700	0,902	0,783	0,754	0,796	0,972	0,861	0,845	0,808	0,924	0,922	0,906	0,796	0,652	0,914	0,571	0,878
	PIC	0,63	0,48	0,70	0,75	0,47	0,71	0,58	0,52	0,58	0,89	0,65	0,64	0,58	0,77	0,78	0,72	0,55	0,45	0,73	0,33	0,70
	PE	0,533	0,127	0,493	0,533	0,224	0,424	0,445	0,205	0,404	0,758	0,404	0,556	0,304	0,629	0,629	0,604	0,252	0,304	0,445	0,103	0,533
	IP	2,11	0,86	1,93	2,11	1,09	1,65	1,73	1,04	1,58	2,24	1,58	2,24	1,28	2,71	2,71	2,53	1,15	1,28	1,73	0,81	2,11
	p HW	0,23459	0,00112	0,55951	0,58931	0,22308	0,55219	0,05292	0,26427	0,28527	0,52619	0,77702	0,64707	0,64313	0,48899	0,0821	0,99545	0,26332	0,15891	0,32928	1	0,06331
GND	N	71	71	71	71	71	71	71	71	70	71	71	71	71	71	71	71	71	71	71	71	71
	2N	142	142	142	142	142	142	142	142	140	142	142	142	142	142	142	142	142	142	142	142	142
	Na	6	5	7	7	6	8	7	6	4	15	5	7	5	9	9	9	5	4	10	6	8
	Ne	3,061	2,160	3,423	4,189	3,063	4,906	2,804	3,281	2,800	9,258	2,817	3,660	2,837	6,052	5,113	6,377	3,106	1,966	7,115	1,733	3,046
	FAM	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,036	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035	0,035
	Ho	0,592	0,549	0,775	0,620	0,620	0,746	0,648	0,676	0,600	0,817	0,535	0,676	0,634	0,775	0,718	0,775	0,690	0,521	0,831	0,437	0,704
	He	0,573	0,537	0,708	0,761	0,673	0,796	0,643	0,695	0,643	0,892	0,645	0,727	0,647	0,835	0,804	0,843	0,678	0,491	0,859	0,423	0,672
	PC	0,151	0,301	0,119	0,091	0,151	0,068	0,226	0,149	0,179	0,032	0,186	0,122	0,177	0,053	0,066	0,052	0,162	0,343	0,043	0,356	0,182
	l emt	6,6	3,3	8,4	11,0	6,6	14,7	4,4	6,7	5,6	31,3	5,4	8,2	5,7	18,7	15,0	19,3	6,2	2,9	23,2	2,8	5,5
	PD	0,849	0,699	0,881	0,909	0,849	0,932	0,774	0,851	0,821	0,968	0,814	0,878	0,823	0,947	0,934	0,948	0,838	0,657	0,957	0,644	0,818
	PIC	0,63	0,47	0,68	0,73	0,62	0,77	0,58	0,67	0,58	0,88	0,58	0,68	0,60	0,81	0,78	0,82	0,62	0,41	0,84	0,39	0,62
	PE	0,281	0,234	0,558	0,315	0,315	0,504	0,352	0,399	0,291	0,631	0,220	0,392	0,340	0,553	0,457	0,553	0,413	0,213	0,658	0,138	0,435
	IP	1,22	1,11	2,25	1,31	1,31	1,97	1,42	1,57	1,25	2,73	1,08	1,54	1,38	2,22	1,78	2,22	1,61	1,06	2,96	0,89	1,69
	p HW	0,26321	0,61057	0,96667	0,42149	0,23035	0,94795	0,01371	0,00921	0,15039	0,10489	0,00869	0,12908	0,79006	0,38548	0,36048	0,13603	0,68201	0,4215	0,48821	0,95016	0,20733
GRC	N	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	50
	2N	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	100
	Na	5	3	7	8	5	7	6	6	4	10	5	6	3	9	5	7	4	2	8	3	5
	Ne	2,338	2,978	3,998	3,413	2,242	4,247	2,535	4,042	2,695	5,459	3,030	3,084	2,760	6,049	4,112	2,034	3,100	1,733	4,077	1,299	3,668
	FAM	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,049	0,050
	Ho	0,608	0,608	0,725	0,529	0,412	0,745	0,529	0,765	0,706	0,706	0,725	0,667	0,608	0,824	0,784	0,373	0,686	0,529	0,725	0,176	0,700
	He	0,572	0,664	0,750	0,707	0,554	0,765	0,606	0,753	0,629	0,817	0,670	0,676	0,638	0,835	0,757	0,508	0,677	0,423	0,755	0,230	0,727
	PC	0,288	0,180	0,106	0,133	0,266	0,103	0,205	0,107	0,216	0,073	0,159	0,158	0,198	0,061	0,103	0,303	0,179	0,470	0,103	0,632	0,135
	l emt	3,5	5,5	9,5	7,5	3,8	9,7	4,9	9,3	4,6	13,6	6,3	6,3	5,0	16,4	9,7	3,3	5,6	2,1	9,7	1,6	7,4
	PD	0,712	0,820	0,894	0,867	0,734	0,897	0,795	0,893	0,784	0,927	0,841	0,842	0,802	0,939	0,897	0,697	0,821	0,530	0,897	0,368	0,865
	PIC	0,50	0,59	0,72	0,66	0,47	0,73	0,54	0,72	0,56	0,79	0,63	0,63	0,58	0,81	0,72	0,47	0,61	0,34	0,72	0,22	0,68
	PE	0,300	0,300	0,477	0,215	0,121	0,501	0,215	0,543	0,437	0,437	0,469	0,379	0,310	0,643	0,570	0,098	0,407	0,223	0,469	0,024	0,428
	IP	1,28	1,28	1,86	1,06	0,85	1,96	1,06	2,17	1,70	1,70	1,82	1,50	1,30	2,83	2,32	0,80	1,59	1,08	1,82	0,61	1,67
	p HW	0,11454	0,60403	0,82227	0,02216	0,09107	0,27299	0,59317	0,45485	0,31747	0,04356	0,86267	0,10379	0,44567	0,6006	0,99615	0,07154	0,76321	0,10485	0,17646	0,03632	0,24643
KIV	N	61	61	61	61	61	61	61	61	61	60	61	61	61	61	60	61	61	61	61	61	60
	2N	122	122	122	122	122	122	122	122	122	120	122	122	122	122	120	122	122	122	122	122	120
	Na	4	5	8	8	5	6	5	5	4	13	4	5	4	9	6	7	5	4	8	4	7
	Ne	2,391	2,556	4,010	5,094	3,627	4,456	3,290	3,850	2,449	7,080	2,052	3,496	3,084	5,108	5,000	4,144	2,156	1,964	4,716	2,128	3,214
	FAM	0,041	0,041	0,041	0,041	0,041	0,041	0,041	0,041	0,041	0,042	0,041	0,041	0,041	0,041	0,042	0,041	0,041	0,041	0,041	0,041	0,042
	Ho	0,623	0,639	0,738	0,689	0,639	0,738	0,639	0,770	0,623	0,850	0,590	0,803	0,623	0,787	0,783	0,803	0,541	0,492	0,902	0,541	0,650
	He	0,582	0,609	0,751	0,804	0,724	0,776	0,696	0,740	0,592	0,859	0,513	0,714	0,676	0,804	0,800	0,759	0,536	0,491	0,788	0,530	0,689
	PC	0,246	0,257	0,098	0,077	0,146	0,087	0,141	0,132	0,261	0,051	0,399	0,166	0,155	0,070	0,084	0,102	0,304	0,309	0,105	0,320	0,141
	l emt	4,1	3,9	10,2	13,1	6,8	11,5	7,1	7,6	3,8	19,6	2,5	6,0	6,4	14,3	11,9	9,8	3,3	3,2	9,5	3,1	7,1
	PD	0,754	0,743	0,902	0,923	0,854	0,913	0,859	0,868	0,739	0,949	0,601	0,834	0,845	0,930	0,916	0,898	0,696	0,691	0,895	0,680	0,859
	PIC	0,52	0,53	0,73	0,78	0,68	0,74	0,65	0,70	0,51	0,84	0,40	0,67	0,63	0,78	0,77	0,72	0,44	0,43	0,76	0,43	0,64
	PE	0,319	0,341	0,496	0,411	0,341	0,489	0,341	0,552	0,319	0,695	0,279	0,605	0,327	0,575	0,568	0,605	0,226	0,188	0,799	0,226	0,355
	IP	1,33	1,39	1,94	1,61	1,39	1,91	1,39	2,21	1,33	3,33	1,22	2,54	1,35	2,35	2,31	2,54	1,09	1,00	5,08	1,09	1,43
	p HW	0,22283	0,0849	0,43528	0,03306	0,00173	0,43271	0,2359	0,18148	0,51532	0,12316	0,2753	0,06074	0,48823	0,35865	0,35352	0,6221	1	0,78256	0,02287	0,61281	0,74612

## ANEXO III

TABELA 22 - Número de alelos por marcador e por população.

MARCADOR	POPULAÇÃO INDÍGENA					POPULAÇÃO BRASILEIRA				
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
(N)	76	71	51	61	60	268	211	182	241	240
D3S1358	5	6	5	4	3	9	12	9	9	8
TH01	7	5	3	5	4	6	8	6	6	8
D21S11	7	7	7	8	9	18	20	15	20	17
D18S51	8	7	8	8	10	15	17	17	13	19
D5S818	5	6	5	5	5	9	8	8	9	9
D13S317	7	8	7	6	6	7	9	7	10	11
D7S820	6	7	6	5	6	8	10	8	8	8
D16S539	6	6	6	5	5	7	7	7	7	8
CSF1PO	5	4	4	4	4	9	8	10	9	10
SE33	16	15	10	13	11	40	41	33	39	45
VWA	5	5	5	4	4	9	10	9	10	9
D8S1179	5	7	6	5	6	9	10	9	10	10
TPOX	4	5	3	4	6	8	7	7	8	7
FGA	8	9	9	9	9	17	21	17	19	18
D2S1338	9	9	5	6	7	11	12	12	13	14
D19S433	9	9	7	7	8	14	12	14	14	16
D10S1248	4	5	4	5	7	9	11	9	10	8
D22S1045	5	4	2	4	3	9	9	10	10	10
D1S1656	11	10	8	8	8	17	16	17	16	16
D2S441	5	6	3	4	4	9	10	9	11	10
D12S391	9	8	5	7	8	19	19	16	17	17
NÚMERO MÉDIO DE ALELOS POR MARCADOR	6,95	7,05	5,62	6,00	6,33	12,33	13,19	11,86	12,76	13,24

## ANEXO IV

**TABELA 23** - Valores de Rst par a par, entre as populações indígenas, para os marcadores *TH01*, *D21S11*, *D18S51*, *D7S820*, *D16S539*, *SE33*, *VWA*, *D2S1338*, *D19S433* e *D10S1248*.

Valores de Rst par a par por população para TH01					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,000				
GRC	0,199	0,249			
KIV	0,297	0,345	0,010		
KRC	0,255	0,303	0,000	0,000	

Valores de Rst par a par por população para D7S820					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,007				
GRC	0,000	0,003			
KIV	0,176	0,245	0,186		
KRC	0,040	0,094	0,045	0,046	

Valores de Rst par a par por população para D21S11					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,087				
GRC	0,258	0,056			
KIV	0,105	0,000	0,040		
KRC	0,149	0,003	0,014	0,000	

Valores de Rst par a par por população para D16S539					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,042				
GRC	0,387	0,199			
KIV	0,050	0,000	0,259		
KRC	0,058	0,000	0,236	0,000	

Valores de Rst par a par por população para D18S51					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,049				
GRC	0,190	0,068			
KIV	0,056	0,218	0,386		
KRC	0,000	0,005	0,080	0,068	

Valores de Rst par a par por população para SE33					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,000				
GRC	0,060	0,037			
KIV	0,133	0,101	0,005		
KRC	0,236	0,199	0,060	0,015	

Valores de Rst par a par por população para Vwa					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,066				
GRC	0,186	0,046			
KIV	0,263	0,084	0,000		
KRC	0,314	0,138	0,004	0,010	

Valores de Rst par a par por população para D2S1338					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,000				
GRC	0,028	0,009			
KIV	0,000	0,000	0,000		
KRC	0,000	0,000	0,021	0,000	

Valores de Rst par a par por população para D19S433					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,000				
GRC	0,166	0,207			
KIV	0,121	0,094	0,465		
KRC	0,089	0,065	0,418	0,000	

Valores de Rst par a par por população para D10S1248					
	GKW	GND	GRC	KIV	KRC
GKW					
GND	0,138				
GRC	0,026	0,042			
KIV	0,160	0,000	0,051		
KRC	0,028	0,025	0,000	0,026	

GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. (+)  $0 < R_{st} < 0,05$ ; (++)  $0,05 < R_{st} < 0,15$ ; (+++)  $0,15 < R_{st} < 0,25$ .

**TABELA 24** - Valores de Rst par a par, entre as populações indígenas e a população brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, vWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 e D12S391.

(continua)

VALORES DE RST POR POPULAÇÃO - LOCO TH01										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,199	0,249								
KIV	0,297	0,345	0,010							
KRC	0,255	0,303	0,000	0,000						
S	0,062	0,090	0,034	0,108	0,075					
SE	0,055	0,084	0,043	0,122	0,087	0,000				
N	0,034	0,060	0,071	0,160	0,122	0,004	0,001			
NE	0,042	0,069	0,060	0,147	0,110	0,001	0,000	0,000		
CO	0,056	0,084	0,040	0,117	0,083	0,000	0,000	0,002	0,000	
VALORES DE RST POR POPULAÇÃO - LOCO D21S11										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,087									
GRC	0,258	0,056								
KIV	0,105	0,000	0,040							
KRC	0,149	0,003	0,014	0,000						
S	0,236	0,038	0,000	0,024	0,004					
SE	0,248	0,045	0,000	0,030	0,008	0,000				
N	0,181	0,014	0,009	0,005	0,000	0,002	0,004			
NE	0,211	0,026	0,002	0,014	0,000	0,000	0,000	0,000		
CO	0,226	0,033	0,000	0,020	0,002	0,000	0,000	0,000	0,000	
VALORES DE RST POR POPULAÇÃO - LOCO D5S818										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,008									
GRC	0,000	0,000								
KIV	0,000	0,000	0,000							
KRC	0,000	0,014	0,000	0,003						
S	0,429	0,495	0,434	0,474	0,389					
SE	0,395	0,461	0,398	0,439	0,353	0,000				
N	0,252	0,324	0,255	0,296	0,214	0,029	0,018			
NE	0,388	0,455	0,392	0,432	0,349	0,000	0,000	0,023		
CO	0,335	0,407	0,339	0,381	0,296	0,005	0,000	0,006	0,003	

**TABELA 24** - Valores de Rst par a par, entre as populações indígenas e a população brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, VWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 E D12S391.

(continuação)

VALORES DE RST POR POPULAÇÃO - LOCO D7S820										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,007									
GRC	0,000	0,003								
KIV	0,176	0,245	0,186							
KRC	0,040	0,094	0,045	0,046						
S	0,106	0,154	0,107	0,000	0,029					
SE	0,000	0,003	0,000	0,000	0,000	0,000				
N	0,119	0,163	0,117	0,000	0,041	0,000	0,000			
NE	0,196	0,249	0,195	0,018	0,093	0,015	0,002	0,003		
CO	0,181	0,232	0,180	0,014	0,083	0,012	0,001	0,002	0,000	
VALORES DE RST POR POPULAÇÃO - LOCO D16S539										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,042									
GRC	0,387	0,199								
KIV	0,050	0,000	0,259							
KRC	0,058	0,000	0,236	0,000						
S	0,366	0,203	0,000	0,227	0,211					
SE	0,280	0,133	0,000	0,150	0,137	0,003				
N	0,135	0,037	0,036	0,042	0,035	0,045	0,018			
NE	0,259	0,115	0,008	0,129	0,116	0,010	0,000	0,010		
CO	0,292	0,139	0,003	0,157	0,143	0,005	0,000	0,018	0,000	
VALORES DE RST POR POPULAÇÃO - LOCO SE33										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,060	0,037								
KIV	0,133	0,101	0,005							
KRC	0,236	0,199	0,060	0,015						
S	0,066	0,043	0,000	0,010	0,068					
SE	0,055	0,034	0,000	0,016	0,079	0,000				
N	0,012	0,001	0,014	0,064	0,151	0,019	0,012			
NE	0,046	0,027	0,000	0,021	0,087	0,000	0,000	0,008		
CO	0,030	0,014	0,001	0,036	0,110	0,005	0,001	0,001	0,000	

**TABELA 24** - Valores de Rst par a par, entre as populações indígenas e a população

brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, VWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 E D12S391.

(continuação)

VALORES DE RST POR POPULAÇÃO - LOCO VWA										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,066									
GRC	0,186	0,046								
KIV	0,263	0,084	0,000							
KRC	0,314	0,138	0,004	0,010						
S	0,094	0,017	0,000	0,000	0,002					
SE	0,089	0,014	0,000	0,000	0,004	0,000				
N	0,135	0,028	0,000	0,000	0,004	0,000	0,000			
NE	0,106	0,021	0,000	0,000	0,001	0,000	0,000	0,000		
CO	0,101	0,019	0,000	0,000	0,001	0,000	0,000	0,000	0,000	

VALORES DE RST POR POPULAÇÃO - LOCO D8S1179										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,020	0,058								
KIV	0,045	0,100	0,000							
KRC	0,000	0,020	0,002	0,023						
S	0,042	0,076	0,000	0,000	0,017					
SE	0,015	0,040	0,000	0,000	0,000	0,003				
N	0,025	0,056	0,000	0,000	0,006	0,000	0,000			
NE	0,003	0,021	0,000	0,010	0,000	0,014	0,001	0,005		
CO	0,038	0,069	0,000	0,000	0,015	0,000	0,003	0,000	0,013	

VALORES DE RST POR POPULAÇÃO - LOCO TPOX										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,000	0,007								
KIV	0,013	0,006	0,041							
KRC	0,025	0,016	0,059	0,000						
S	0,172	0,155	0,234	0,076	0,056					
SE	0,227	0,209	0,293	0,117	0,094	0,003				
N	0,153	0,140	0,182	0,088	0,074	0,021	0,007			
NE	0,200	0,183	0,248	0,109	0,090	0,011	0,000	0,001		
CO	0,182	0,168	0,215	0,110	0,094	0,029	0,013	0,000	0,004	

**TABELA 24** - Valores de Rst par a par, entre as populações indígenas e a população

brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, VWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 E D12S391.

(continuação)

VALORES DE RST POR POPULAÇÃO - LOCO D19S433										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,166	0,207								
KIV	0,121	0,094	0,465							
KRC	0,089	0,065	0,418	0,000						
S	0,056	0,085	0,043	0,343	0,298					
SE	0,022	0,041	0,069	0,260	0,219	0,003				
N	0,005	0,017	0,102	0,206	0,167	0,018	0,002			
NE	0,004	0,016	0,096	0,199	0,162	0,019	0,003	0,000		
CO	0,034	0,056	0,059	0,291	0,248	0,000	0,000	0,006	0,007	
VALORES DE RST POR POPULAÇÃO - LOCO D22S1045										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,004									
GRC	0,026	0,000								
KIV	0,113	0,070	0,050							
KRC	0,000	0,000	0,014	0,095						
S	0,045	0,028	0,017	0,000	0,033					
SE	0,028	0,013	0,005	0,000	0,018	0,000				
N	0,014	0,002	0,000	0,004	0,006	0,008	0,001			
NE	0,044	0,027	0,016	0,000	0,032	0,000	0,000	0,008		
CO	0,015	0,004	0,000	0,000	0,007	0,005	0,000	0,000	0,006	
VALORES DE RST POR POPULAÇÃO - LOCO D1S1656										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,000									
GRC	0,000	0,000								
KIV	0,016	0,045	0,021							
KRC	0,000	0,013	0,000	0,000						
S	0,036	0,009	0,025	0,115	0,064					
SE	0,036	0,009	0,025	0,114	0,064	0,000				
N	0,015	0,000	0,007	0,078	0,036	0,001	0,001			
NE	0,042	0,013	0,031	0,125	0,072	0,000	0,000	0,003		
CO	0,027	0,004	0,018	0,100	0,053	0,000	0,000	0,000	0,000	



**TABELA 24** - Valores de Rst par a par, entre as populações indígenas e a população brasileira, para os marcadores TH01, D21S11, D5S818, D7S820, D16S539, SE33, VWA, D8S1179, TPOX, D19S433, D22S1045, D1S1656, D2S441 E D12S391.

(conclusão)

VALORES DE RST POR POPULAÇÃO - LOCO D2S441										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,004									
GRC	0,000	0,001								
KIV	0,132	0,000	0,123							
KRC	0,152	0,000	0,140	0,000						
S	0,058	0,044	0,053	0,049	0,048					
SE	0,057	0,041	0,051	0,047	0,046	0,000				
N	0,039	0,022	0,034	0,029	0,028	0,008	0,003			
NE	0,054	0,039	0,049	0,045	0,043	0,000	0,000	0,003		
CO	0,064	0,049	0,058	0,054	0,053	0,000	0,000	0,011	0,000	

VALORES DE RST POR POPULAÇÃO - LOCO D12S391										
	GKW	GND	GRC	KIV	KRC	S	SE	N	NE	CO
GKW										
GND	0,011									
GRC	0,006	0,000								
KIV	0,019	0,033	0,014							
KRC	0,000	0,022	0,016	0,027						
S	0,017	0,031	0,027	0,034	0,004					
SE	0,024	0,039	0,035	0,042	0,008	0,000				
N	0,003	0,016	0,013	0,019	0,000	0,003	0,008			
NE	0,007	0,020	0,017	0,023	0,000	0,001	0,004	0,000		
CO	0,013	0,026	0,023	0,029	0,001	0,000	0,000	0,001	0,000	

## ANEXO V

**TABELA 25** - Resultados de Fst par a par entre Guaranis, Kaingangs, ameríndios (América do Norte, América Central e América do Sul) e Parentais (Japão, África do Sul, Portugal, Alemanha e Espanha).

(continua)

	2	3	4	5	6	7	8	9	10	11	12
1	0.009	0.095	0.091	0.087	0.112	0.168	0.129	0.066	0.046	0.078	0.073
2		0.065	0.063	0.057	0.070	0.145	0.111	0.047	0.023	0.048	0.034
3			0.071	0.069	0.065	0.129	0.093	0.060	0.055	0.043	0.045
4				-0.000	0.030	0.093	0.079	0.029	0.061	0.029	0.041
5					0.045	0.105	0.088	0.039	0.065	0.047	0.056
6						0.052	0.039	0.022	0.066	0.018	0.031
7							0.012	0.058	0.132	0.070	0.101
8								0.037	0.106	0.051	0.076
9									0.053	0.022	0.031
10										0.028	0.008
11											0.009
	13	14	15	16	17	18	19	20	21	22	23
1	0.060	0.052	0.070	0.109	0.083	0.082	0.085	0.090	0.176	0.072	0.107
2	0.031	0.022	0.042	0.072	0.048	0.048	0.060	0.065	0.150	0.039	0.082
3	0.051	0.008	0.054	0.062	0.020	0.039	0.027	0.073	0.111	0.029	0.102
4	0.037	0.052	0.044	0.040	0.021	0.011	0.031	0.056	0.093	0.008	0.069
5	0.053	0.048	0.064	0.043	0.029	0.019	0.041	0.073	0.114	0.017	0.083
6	0.034	0.052	0.042	0.026	0.012	-0.002	0.007	0.035	0.050	0.024	0.043
7	0.086	0.123	0.097	0.077	0.066	0.030	0.055	0.073	0.008	0.084	0.070
8	0.071	0.084	0.075	0.058	0.044	0.018	0.032	0.056	0.020	0.061	0.068
9	0.021	0.037	0.025	0.026	0.016	0.010	0.012	0.032	0.058	0.022	0.041
10	0.002	0.036	0.016	0.041	0.031	0.034	0.029	0.022	0.144	0.020	0.048
11	-0.002	0.032	0.016	0.027	-0.009	-0.001	0.002	0.033	0.070	-0.007	0.039
12	-0.008	0.025	-0.007	0.028	0.011	0.015	0.006	0.011	0.106	0.004	0.047
13		0.027	-0.008	0.029	0.004	0.013	0.010	0.016	0.084	-0.004	0.036
14			0.034	0.057	0.019	0.031	0.035	0.065	0.108	0.018	0.095
15				0.039	0.020	0.022	0.012	0.017	0.095	0.011	0.048
16					0.011	0.010	0.002	0.016	0.089	0.018	0.010
17						-0.003	-0.005	0.030	0.060	-0.008	0.041
18							-0.007	0.008	0.044	-0.004	0.019
19								0.003	0.064	-0.004	0.018
20									0.092	0.027	0.009
21										0.084	0.091
22											0.048
	24	25	26	27	28	29	30	31	32	33	34
1	0.091	0.081	0.107	0.099	0.094	0.101	0.071	0.057	0.073	0.119	0.057
2	0.060	0.067	0.074	0.065	0.065	0.074	0.040	0.035	0.052	0.068	0.025
3	0.060	0.043	0.098	0.089	0.040	0.083	0.070	0.036	0.101	0.120	0.065

**TABELA 25** - Resultados de Fst par a par entre Guaranis, Kaingangs, ameríndios (América do

Norte, América Central e América do Sul) e parentais (Japão, África do Sul, Portugal, Alemanha e Espanha).

(continuação)

4	0.037	0.054	0.073	0.057	0.024	0.073	0.036	0.015	0.059	0.053	0.029
5	0.042	0.053	0.079	0.073	0.038	0.087	0.040	0.029	0.068	0.048	0.034
6	0.017	0.033	0.064	0.043	0.015	0.047	0.027	-0.001	0.075	0.072	0.029
7	0.056	0.049	0.118	0.102	0.081	0.091	0.056	0.044	0.099	0.142	0.079
8	0.041	0.025	0.093	0.080	0.059	0.066	0.055	0.017	0.087	0.128	0.060
9	0.020	0.018	0.059	0.028	0.013	0.040	0.020	-0.012	0.026	0.064	0.024
10	0.046	0.059	0.031	0.035	0.041	0.041	0.030	0.027	0.058	0.066	0.019
11	0.024	0.040	0.045	0.040	0.011	0.039	0.023	0.003	0.060	0.070	0.027
12	0.032	0.047	0.024	0.012	0.014	0.028	0.014	0.001	0.038	0.039	-0.002
13	0.028	0.045	0.033	0.021	0.014	0.034	0.005	0.005	0.026	0.048	0.015
14	0.057	0.038	0.090	0.072	0.038	0.072	0.044	0.025	0.069	0.074	0.035
15	0.038	0.045	0.037	0.017	0.022	0.039	0.020	0.005	0.027	0.060	0.012
16	0.003	0.029	0.030	0.021	0.001	0.023	0.028	-0.000	0.067	0.056	0.029
17	0.012	0.024	0.045	0.037	0.001	0.029	0.021	-0.001	0.053	0.054	0.029
18	-0.004	0.008	0.034	0.026	0.005	0.030	0.007	-0.018	0.043	0.059	0.004
19	-0.002	0.005	0.028	0.014	-0.008	0.020	0.019	-0.015	0.050	0.080	0.016
20	0.015	0.028	0.019	0.006	0.017	0.006	0.024	-0.005	0.043	0.074	0.010
21	0.065	0.055	0.143	0.117	0.077	0.111	0.071	0.049	0.104	0.151	0.095
22	0.019	0.035	0.034	0.032	0.005	0.044	0.018	0.003	0.041	0.047	0.009
23	0.002	0.043	0.030	0.025	0.017	0.022	0.033	0.009	0.062	0.098	0.043
24		0.022	0.030	0.024	0.006	0.027	0.021	-0.011	0.043	0.072	0.024
25			0.072	0.051	0.023	0.041	0.034	-0.001	0.061	0.100	0.035
26				0.017	0.036	0.032	0.046	0.020	0.059	0.064	0.023
27					0.020	0.021	0.034	0.000	0.024	0.059	0.012
28						0.024	0.029	-0.014	0.055	0.059	0.027
29							0.045	0.006	0.058	0.083	0.036
30								0.007	0.032	0.048	0.013
31									0.021	0.052	-0.002
32										0.074	0.026
33											0.015
	35	36	37	38	39	40	41	42	43	44	45
1	0.048	0.067	0.097	0.051	0.093	0.075	0.086	0.072	0.054	0.200	0.211
2	0.011	0.032	0.069	0.034	0.065	0.045	0.050	0.043	0.062	0.198	0.202
3	0.036	0.055	0.117	0.019	0.096	0.054	0.059	0.049	0.074	0.219	0.213
4	0.021	0.012	0.039	0.033	0.043	0.040	0.048	0.018	0.076	0.200	0.185
5	0.028	0.020	0.035	0.039	0.046	0.041	0.048	0.030	0.084	0.216	0.198
6	0.024	0.022	0.055	0.009	0.035	0.015	0.034	0.000	0.071	0.183	0.166
7	0.076	0.042	0.094	0.042	0.062	0.052	0.088	0.050	0.117	0.237	0.209
8	0.056	0.050	0.066	0.019	0.049	0.031	0.071	0.039	0.083	0.222	0.201
9	0.010	0.020	0.028	-0.006	0.003	-0.008	0.023	-0.007	0.032	0.169	0.160
10	0.020	0.021	0.087	0.042	0.064	0.046	0.075	0.042	0.054	0.193	0.192
11	0.021	0.008	0.051	0.004	0.043	0.026	0.058	0.014	0.049	0.153	0.145

**TABELA 25** - Resultados de Fst par a par entre Guaranis, Kaingangs, ameríndios (América do Norte, América Central e América do Sul) e parentais (Japão, África do Sul,

Portugal, Alemanha e Espanha).

(continuação)

12	-0.007	0.012	0.049	0.018	0.035	0.026	0.041	0.004	0.037	0.182	0.175
13	-0.001	-0.002	0.055	0.017	0.035	0.022	0.049	0.012	0.039	0.154	0.149
14	-0.002	0.035	0.068	0.008	0.063	0.033	0.029	0.032	0.059	0.197	0.198
15	0.001	0.015	0.057	0.024	0.036	0.024	0.051	0.013	0.039	0.182	0.177
16	0.027	0.029	0.053	0.020	0.022	-0.002	0.048	0.007	0.045	0.195	0.173
17	0.015	0.010	0.057	-0.005	0.042	0.013	0.033	0.008	0.046	0.162	0.153
18	0.001	-0.007	0.046	-0.010	0.021	0.002	0.016	-0.016	0.065	0.182	0.161
19	0.019	0.019	0.055	-0.001	0.037	0.006	0.049	-0.001	0.038	0.171	0.155
20	0.019	0.020	0.075	0.025	0.029	0.013	0.054	0.006	0.042	0.197	0.179
21	0.070	0.055	0.114	0.043	0.077	0.060	0.080	0.058	0.124	0.231	0.208
22	0.003	-0.004	0.037	0.014	0.042	0.026	0.056	0.018	0.050	0.161	0.150
23	0.047	0.037	0.094	0.037	0.042	0.013	0.058	0.010	0.053	0.193	0.168
24	0.026	0.021	0.065	0.014	0.029	0.004	0.034	-0.001	0.054	0.184	0.165
25	0.023	0.038	0.064	-0.005	0.034	0.004	0.029	0.008	0.047	0.193	0.177
26	0.049	0.037	0.076	0.063	0.049	0.042	0.094	0.037	0.054	0.214	0.198
27	0.024	0.034	0.059	0.040	0.021	0.012	0.060	0.007	0.043	0.211	0.197
28	0.016	0.027	0.053	0.004	0.023	0.003	0.030	-0.011	0.030	0.165	0.150
29	0.042	0.040	0.078	0.027	0.036	0.018	0.058	0.018	0.041	0.201	0.186
30	0.000	-0.003	0.045	0.014	0.021	0.012	0.032	0.003	0.056	0.168	0.153
31	-0.004	0.006	0.023	-0.019	-0.006	-0.018	0.011	-0.030	0.010	0.162	0.151
32	0.017	0.023	0.065	0.047	0.032	0.032	0.048	0.024	0.066	0.211	0.204
33	0.022	0.027	0.042	0.075	0.030	0.050	0.076	0.047	0.081	0.239	0.224
34	-0.013	-0.007	0.026	0.024	0.011	0.017	0.047	0.007	0.036	0.205	0.190
35		-0.008	0.030	0.003	0.012	0.002	-0.000	-0.007	0.041	0.183	0.174
36			0.037	0.016	0.019	0.020	0.049	0.011	0.057	0.186	0.169
37				0.043	0.005	0.027	0.077	0.034	0.048	0.219	0.205
38					0.011	-0.014	-0.001	-0.021	0.026	0.150	0.142
39						-0.023	0.032	-0.013	0.030	0.193	0.177
40							0.003	-0.022	0.025	0.176	0.161
41								-0.017	0.073	0.201	0.191
42									0.025	0.164	0.151
43										0.181	0.175
44											0.061
	46	47	48								
1	0.195	0.190	0.193								
2	0.187	0.182	0.185								
3	0.196	0.177	0.183								
4	0.189	0.169	0.174								
5	0.204	0.186	0.192								
6	0.171	0.164	0.168								
7	0.215	0.211	0.214								
8	0.201	0.195	0.199								

**TABELA 25** - Resultados de Fst par a par entre Guaranis, Kaingangs, ameríndios (América do Norte, América Central e América do Sul) e parentais (Japão, África do Sul, Portugal, Alemanha e Espanha).

(conclusão)

9	0.161	0.151	0.155
10	0.173	0.165	0.167
11	0.129	0.115	0.119
12	0.163	0.156	0.158
13	0.135	0.126	0.129
14	0.181	0.170	0.175
15	0.165	0.155	0.158
16	0.177	0.166	0.169
17	0.137	0.121	0.125
18	0.160	0.147	0.151
19	0.155	0.142	0.146
20	0.181	0.175	0.177
21	0.209	0.202	0.207
22	0.145	0.129	0.134
23	0.171	0.165	0.167
24	0.167	0.158	0.161
25	0.180	0.171	0.176
26	0.193	0.184	0.185
27	0.198	0.191	0.192
28	0.154	0.140	0.144
29	0.184	0.178	0.178
30	0.155	0.151	0.153
31	0.152	0.142	0.145
32	0.207	0.199	0.202
33	0.228	0.221	0.222
34	0.195	0.190	0.193
35	0.173	0.167	0.171
36	0.169	0.158	0.161
37	0.212	0.204	0.206
38	0.136	0.127	0.131
39	0.189	0.184	0.185
40	0.165	0.159	0.161
41	0.189	0.184	0.187
42	0.155	0.148	0.150
43	0.169	0.162	0.164
44	0.043	0.053	0.052
45	0.070	0.059	0.066
46		0.019	0.016
47			0.001

1=GKW, 2=GND, 3=GRC, 4=KIV, 5=KRC, 6=Athabaskan, 7=Inupiat, 8=Yupik, 9=Amerindios Bolivia, 10=El Conchagua, 11= San Alejo, 12=Panchimalco, 13=Izalco, 14=-Mbyá-Guarani (Argentina), 15= Maya (Guatemala), 16=Apache, 17= Cherokee, 18=Chippewa, 19=Cochimi, 20=Cora, 21=Eskimo, 22=Miwok, 23=Seri, 24=Yavapai, 25=Tarahumara, 26=Huichol, 27=Tepehuano, 28=Mayos, 29=Mexicaneros, 30=Terena, 31=Quechua, 32=Emerillon, 33=Palikut, 34=Wayampi, 35=Kalina, 36=Apalai, 37=Matsiguenga, 38=Pilaga, 39=Toba\_Cho, 40=Toba\_Fo, 41=Wichi-Chi, 42=Wichi-for, 43=Kairitiana e Surui.

## ANEXO VI

**TABELA 26** - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS GUARANI E KAINGANG.

K	REPS	MEAN LNP(K)	STDEV LNP(K)	LN'(K)	LN''(K)	DELTA K
1	3	-20317.933333	0.208167	—	—	—
2	3	-19526.400000	1.345362	791.533333	322.966667	<b>240.059233</b>
3	3	-19057.833333	1.814754	468.566667	327.000000	180.189677
4	3	-18916.266667	34.213643	141.566667	31.433333	0.918737
5	3	-18806.133333	50.543480	110.133333	—	—

**TABELA 27** - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS BRASILEIROS (GUARANI E KAINGANG) E A POPULAÇÃO BRASILEIRA MISCIGENADA.

K	REPS	MEAN LNP(K)	STDEV LNP(K)	LN'(K)	LN''(K)	DELTA K
1	3	-114670.733333	0.461880	—	—	—
2	3	-111681.633333	4.842864	2989.100000	3311.800000	<b>683.851516</b>
3	3	-112004.333333	31.666123	-322.700000	860.833333	27.184677
4	3	-113187.866667	280.071818	-1183.533333	207.966667	0.742548
5	3	-114163.433333	851.926959	-975.566667	94.966667	0.111473
6	3	-115044.033333	1115.145113	-880.600000	3873.433333	3.473479
7	3	-119798.066667	1166.891620	-4754.033333	7271.100000	6.231170
8	3	-117281.000000	2679.851496	2517.066667	5554.833333	2.072814
9	3	-120318.766667	5337.249279	-3037.766667	2264.733333	0.424326
10	3	-121091.800000	5352.150515	-773.033333	—	—

**TABELA 28** - VALORES DE DELTA K GERADOS PELO STRUCTURE HARVESTER PARA OS INDÍGENAS BRASILEIROS (GUARANI E KAINGANG) E AS POPULAÇÕES PARENTAIS..

K	REPS	MEAN LNP(K)	STDEV LNP(K)	LN'(K)	LN''(K)	DELTA K
1	3	-107804.700000	0.264575	—	—	—
2	3	-104462.033333	2.454248	3342.666667	2325.200000	<b>947.418510</b>
3	3	-103444.566667	4.718404	1017.466667	781.733333	165.677501
4	3	-103208.833333	26.094891	235.733333	84.033333	3.220298
5	3	-103057.133333	298.556131	151.700000	61.733333	0.206773
6	3	-102843.700000	280.519304	213.433333	81.300000	0.289820
7	3	-102711.566667	26.643073	132.133333	96.266667	3.613197
8	3	-102675.700000	183.789445	35.866667	7.233333	0.039357
9	3	-102647.066667	149.807221	28.633333	—	—